INSTITUTE *of* STATISTICS
Oxford

Bulletin Vol. 22 *November 1960* *No. 4*

Since it first appeared, a mere cyclostyled broadsheet, in 1939–40, *the* BULLETIN *has been under the editorship and guidance of David Worswick. We must salute, with respect and gratitude, the ending of this long, happy and successful association.*

THE ECONOMICS OF ROADS

INTRODUCTION

By D. L. MUNBY

The following papers raise a number of questions about the economics of roads. They do not raise all the questions that need to be discussed, nor do they provide solutions to the large questions of public policy about roads that remain unsettled. But it can be claimed that, if these questions raised here are not answered, policy cannot be properly based. There has been too little attention to the economics of roads, and too much of the current discussion is based on slogans and biassed by vested interests. Far more searching questions need to be asked. Among them will be those asked here.

Because these questions have not received a great deal of attention in this country, it is not surprising that, when they are raised, they provoke further questions and deeper analysis, which may seem to lose itself in the refined stratosphere of welfare economics in heights from which, as it seems to many of the most subtle minds, there is no return to earth. It may be that some readers will feel that, if nothing can be said about the economics of roads until we have explored all the byways of the theory of consumers' surplus, nothing is gained by raising any economic questions or asking advice of economists. This would be misleading; it would exaggerate the extent of disagreement on these issues among economists, and give too much importance to discussion of refined details of analysis.

All but one of the authors of the papers in this BULLETIN took part in a small private conference at Nuffield College in March 1960. At this were present academic economists from Oxford and elsewhere, as well as economists and others engaged in more official and practical fields. At this conference the first draft of the opening paper was circulated and discussed. What emerged was, in effect, a remarkable measure of agreement about the questions to be asked and the problems to be settled. After this conference, some of the participants produced the other papers, taking up various aspects of the questions raised. Some of the papers have been seen by writers of some others, but it has not, unfortunately, proved possible for all the drafts to be circulated to everybody, either for comment or for revision of their own contributions. Hence, though all the papers have a common origin and a common concern, they take up rather different aspects of these common problems, and treat them rather differently.

It is hoped that the questions raised may provoke further discussion and that others may take them up in criticism, elaboration, or application. Perhaps it may be worth while to outline some of these key questions?

1. What are the economic effects on road and rail competition and on the utilisation of our road system of the existing methods of charging for the use of the roads by the various taxes actually collected from road users?

2. Is it practicable to make better use of the pricing system to achieve a more sensible utilisation of the road system we inherit?

3. What would be the principles of such improved utilisation, if we could, in practice, adopt an optimum charging system?

4. How far can the road system and the use of roads be treated as a number of separate units and problems, each perhaps requiring different treatment?

5. If so, should we distinguish between under-utilised and congested roads, and between the problem of inter-city roads and the urban network, and treat the parking, delivering and access problem separately?

6. What are the economics of roads in cities?

7. How can we decide how much to invest in new roads, both in total and in particular schemes? Can this be done if the roads are not charged for? Is it possible to estimate what people would be ready to pay, if they are not, in fact, asked to pay? If not, how do we avoid over-investment in roads?

8. How much do we really know about our road system and the uses to which it is put?

9. What further information is required, as a basis for a more rational and comprehensive policy?

THE ROADS AS ECONOMIC ASSETS

By D. L. Munby

1. Introduction

It is often thought that the major problem of transport policy lies with the railways. By and large, since the 1953 Act, this is not so. Granted the completion of the move towards commercial freedom which was initiated and largely fulfilled in the Act, and granted some rearrangement of their capital structure, the railways should be in a position to follow a rational economic policy. In so far as they remain a problem, it is either because this policy is not logically followed out, or because they are prevented from enjoying its benefits by the fact that road users are not charged for their use of the roads on any coherent principles. There is little virtue in being rational in an irrational world; some degree of general irrationality would then suit everybody better. But this paper is not concerned with a second-best solution, whereby the suitable amount of irrationality should be discovered for the railways, given the other irrationalities in the situation. It is the contention that the real problem is that of the roads, and that little has been done to think out what would be involved in a rational policy for the roads, as part of a total transport policy. Even if a rational policy proves in the last resort impossible because of institutional factors which prevent an adequate adjustment, much will be gained from elaborating what would be involved in such a policy if it were possible. This should provide a more solid basis for a second-best policy than is afforded by deliberate avoidance of thought about the fundamentals.

By a rational policy, in a situation where we have competition between two alternative methods of providing the same services, each with different cost structures, is meant a policy which relates charges to costs in such a way that each provider can (in the long run) cover his long-run marginal costs (including his indivisible costs),[1] at the same time that any particular service is charged no more than the marginal divisible cost, where this is important; in the context, the latter qualification means that a user should not be tempted away from the provider with the lower divisible cost, and should not be precluded from using the service where he is ready to pay the divisible cost. The detailed analysis of these cases has been provided by Arthur Lewis in his article on ‘ Fixed Costs ’.[2] It is not assumed that it is possible precisely to fulfil all these conditions, or that too much precision is desirable. It is however relevant to discover exactly how important, for example, is the divisible cost criterion in practice; equally, we may need to know whether it is possible to cover indivisible costs, and whether the special conditions (discussed by

---

[1] Indivisible costs are often labour costs in transport, so that, though escapable in terms of time without much delay, they are nevertheless necessarily incurred if any service is to be provided at all.

[2] *Overhead Costs* (1949), chapter 1.

Lewis) may not arise where it is necessary forcibly to eliminate one provider of the service in question.

There are two alternative solutions of the road/rail problem, which do not seem acceptable. The first is a proposal sometimes put forward entirely to separate track cost from direct costs and for the government to take over the railway track and run it as it does the roads. The extreme view holds that railway track should be paid for out of taxes, though perhaps admitting that railways might pay petrol tax on their consumption of diesel oil. (If the government is meant to collect the cost of the track from railway users, they are in no better position than British Railways, unless it can be shown that there is some economy to be gained from divorce of the two parts of the business, which hardly seems probable). The obvious objection to this solution is that, if two industries provide the same service by alternative methods, but with unequal relative proportions of indivisible to total cost, there will be a waste of resources in the industry with the higher proportion of indivisible costs, if charges are merely based on divisible costs. It is not surprising that some railway enthusiasts favour such a solution, as it is likely to favour railways as against roads under most conditions that can be expected. (The solution would of course be acceptable, if at likely levels of operation there was little difference in the proportions of indivisible to divisible cost; but then there would be no problem.) This difficulty can only be met if users have to pay indivisible costs as well as divisible costs.

The second solution is that proposed by Sargent in his *British Transport Policy*, whereby we take as a datum the system of charging for the roads by means of a petrol tax, and then adjust rail charges so as to ensure that the indivisible costs of the railway system are borne in the same sort of way. This is a possible solution, if it is taken for granted that average road costs are never *below* direct costs, though it seems rather clumsy, besides being difficult to operate. It is not clear why we must regard the present system of road taxation as unalterable. If we could devise a system which did not average out the indivisible costs of the roads, but allowed the cost of use of the roads to be related to direct costs, so that, where these were below average, the road would be able to compete on fair terms with the lowest charges the railways can make, and, where they were above, the railways would not be unfairly penalised, then there would be no need for Sargent's make-do-and-mend, which ignores the wide variation in direct costs of roads of different sorts.

In fact there is every reason to believe that direct road costs are often much *above* the charge made by the petrol tax.

What follows is an attempt to look at these various situations and see whether some greater approximation to the railway system of allocating indivisible costs could not be devised for the roads. This requires primarily an attempt to sort out the various different situations, and in particular the variations in the value of land and the utilisation of road-space in different parts of the country. Most discussion about road policy has been bedevilled by general arguments relevant to little in particular. It is easy to show that the sums collected from road users in total are in excess of the expenditure on roads on both current and capital account, though the rising expenditure on new roads plus a proper allocation of costs such as that of the police might in a few years alter the picture. But it is not clear what economic significance such statements have. If we could show that all the coal industries of the world taken together made large profits, it would not justify the National Coal Board in making large losses, nor would it throw any light on the charging policy that the N.C.B. should follow.

What follows is an attempt to work out the principles of pricing policy for different road situations, supposing that it would be possible to approximate to such a policy. Planning solutions may be better (this is taken up in Section 8); but to discuss the ideal pricing policy may help to bring out clearly the inherent problems. At the same time, an assumption is made that travellers and shippers should, if possible, have the right to choose the method of transport. The justification of this is that they alone can know the total costs they have to incur, as Sargent clearly shows.[1] Another assumption is implicit in what follows, namely that there are no special external economies involved in the use of roads, which cannot (in theory) be collected from the users of the road; this is taken up in Section 9. Clearly, if it can be shown that special external economies are involved in the provision and use of roads, there is a case for charging users less than the cost of providing them.

If we try to sort out the different elements in the road situation, we are faced with problems of considerable complexity, each of which requires different treatment. But there are similar situations in other fields of economic policy, at least in certain cases. The first broad distinction is between new roads and old roads. I start then with the problem of charging for new roads, which are assumed to be built so as to pay for themselves without becoming congested, and then turn to congested roads and under-utilised roads.[2] In each case it is (at least initially) assumed that the road is merely a road between A and B, not used for any other purpose than that of movement, and used continuously during a period.

2. Cost of a New Road

We may take it that the costs of a road are largely capital costs, but that some current costs are involved. The capital cost of a new road includes the cost of the land required and the cost of construction. These should normally be treated as capital items in the balance sheet, and only the interest and depreciation charges set down to current account.

How should these be assessed? At the very least roads should be charged the going rate at which the government borrows. But, as with the

[1] *British Transport Policy,* chapter 1. See also A. M. Milne, *The Economics of Inland Transport,* chapter 4.

[2] Of course these issues are interrelated. If congestion is cured, it may cure under-utilisation. This paper only takes the first step of separating out the different pure cases, though there are some hints about complications.

nationalised industries in general, there is a strong case for arguing that this does not adequately represent the value of capital to the community, and that the public sector should expect to make at the margin a return no less than that expected in the private sector.[1] We need not go here into what that involves.

It is much more difficult to make a proper allowance for depreciation. Roads can almost be regarded as everlasting assets, at least as regards the cost of land, and even the construction cost could on a technical basis be safely amortised over a very long period. Nor is there a great need to fear obsolescence, though one should perhaps be cautious here. It is, after all, not impossible that the rapid growth and expansion of the motor-car will be followed by an equally rapid decline as some other as yet unknown mode of travel supersedes it, which does not require the same sort of track. For our purposes here, however, we do not need to specify any particular period for depreciation.

Maintenance costs may not be great, but cannot be ignored. They are of two kinds. Roads wear out as a result of weather conditions and through the mere passage of time. A part of the maintenance that is required will thus be entirely unrelated to use of the road, and can be broadly related to time. A second part of the cost of maintenance is incurred because of the use of the road. In so far as this element is important, it is important to discover how wear and tear is produced by use. We want to know what are the factors in vehicles or in the way they are driven which cause these costs to be incurred, so that we can make these users pay. We need to know certain technical factors about vehicles, as well as social facts about the way they are used, so that we can rightly allocate the costs incurred by speed, weight, and type of vehicle.

Roads also involve lighting costs, and what may be called signalling costs. This latter element, the cost of the police involved in traffic direction and accident control, should be paid for by the road user, and is a normal current cost of road traffic. We need information about this aspect of the roads.

Lastly, there is the administrative cost of the roads as born by the various highway departments concerned with different aspects of the roads.

To sum up, we have, in the case of a new road, capital costs that are incurred once and for all, current costs which are independent of utilisation of the roads, but must be incurred if the roads are kept open at all (time-maintenance, lighting, signalling, administration), and current costs which vary with utilisation. It may be agreed that the last element is very small in relation to the rest. The other costs whether capital or current can be expressed as a cost per year of a given road (given our assessment of a proper rate of interest and period for depreciation). They have to be collected somehow or other from the users in that period. (The principle would not be very different, though more complicated, if one were to assume that utilisation increased over time, so that some of these costs had to be capitalised and borne by later users. One would have to assume that it had been right to build the road at that time and not later, taking into account the agio of expected returns, the time-shape of costs, and the rate of interest. We ignore these complications by assuming both costs and traffic to be constant over time. It is to be noted in particular that the alternative use cost of land may rise over time.)

---

[1] See 'The Finance of the Nationalised Industries' (BULLETIN OF THE INSTITUTE OF STATISTICS, May 1959), and 'Investing in Coal' (*Oxford Economic Papers*, Oct. 1959); also 'The Cost of Financing the Nationalised Industries' by C. D. Foster (BULLETIN, May 1960).

How then are we to charge road users so as to cover these costs? Our ideal aim should be to make them pay for the total costs incurred, without reducing their use of the road below that level at which they are ready to pay for the utilisation costs they incur. The running charge should ideally be kept down to this level and the other costs have to be charged to the users in some way or other. Briefly, there are three possibilities.

(1) Average cost is charged to everyone. This itself depends on the degree of utilisation achieved, and will require a degree of estimation and guesswork to achieve. The objection to this method of charging is that utilisation is kept below what it would otherwise be, to an extent depending on the elasticity of demand for the services of the road. The excluded users, who would be ready to pay the use-costs they incur but not their proportionate share of the other costs, would not create any costs by being allowed to use the road and would themselves gain. Any contribution they may make to the other costs would mean less for the rest to pay. There is therefore a loss to everybody in this method of charging. (Strictly speaking, on our assumptions, with no failures in forecasting, utilisation will be optimal, and one should say that investment will be lower than with other pricing policies, and that investment which could pay for itself will not be undertaken.)

(2) The other extreme method of charging is to make everyone pay a contribution according to the benefit they gain from the road. In the extreme case his consumer surplus is extracted from everybody. If this policy were generally followed, investment would be larger than it would be otherwise, but, even then, it will only be at the margin that the contribution is not in excess of the costs involved. If our aim is only to collect a contribution of the same size as the total costs, we will have to scale down the charge below the maximum that can be extracted. Once we have decided to scale down the contribution below what people are prepared to pay, we have introduced an element of arbitrariness, and are involved in our third method of charging.

(3) The third method can be characterised as a two-part tariff, though there can be more than two-parts if we wish it. Essentially, it is some form of compromise between the first two forms of charging. If the running charge is the most important, and the other part of the tariff unimportant, it approximates to the average cost solution. If the prior charge element is high, and the running charge only equals the use-cost, then it approximates to the preceding case. Clearly, there is a wide choice open as to how the prior charge is to be levied.

3. CONGESTED ROADS

If an asset is used to capacity, and there is no immediate question of expansion, the most suitable form of pricing is to follow supply and demand and charge such a price as will clear the market. The only alternative is to ration by some method or other, but this will normally be clumsy and difficult, unless there is an obvious criterion available by which to choose the lucky gainers. A price which clears the market sorts out those who really want to use the asset most (given the normal assumptions about a fair distribution of income). It can also be regarded as the marginal cost price. (Graphically the cost curve rises vertically at the full capacity point, but this is a more than formal point). The cost of A having the use of a service, which cannot be immediately, expanded is that B is deprived of its use; the cost to A is the loss of value to B, and this is the marginal cost of the service.[1]

This criterion is the right one for a service which is at least temporarily limited in some absolute way, as for example in the case of the supply of electricity or the capacity of a bus, given certain restrictions on standing. In these cases capacity is a definite fixed limit to output.

This is not the case with roads, as with many other assets. Normally, output can be expanded at increasing cost, even at a ' full capacity point '. The difficulty in the case of the roads arises from the nature of the increasing costs and the flexibility of what may be regarded as full capacity.

One may measure the use of a road by the flow of vehicles past a point per period of time. The flow is clearly an increasing function of the speed of vehicles and a decreasing function of the distance between vehicles. As the speed of vehicles increases, the minimum distance between vehicles will tend to rise, given the normal pattern of driving. (No doubt flow is also affected by the distribution of cars driving at different speeds; for our purposes we can consider the matter in terms of a flow of vehicles moving at uniform speed, though a more realistic analysis might take the form of assuming a normal distribution around some more or less conventional speed.) In terms of a flow of vehicles moving at uniform speed, clearly the flow will be greater at any speed as cars follow each other more closely up to a certain point. If we start with a flow of cars travelling at a given speed, and if more cars join the flow, at some critical point (and this will be lower the higher is the initial speed) speed will decrease as cars crowd closer together and flow increases. If the process of crowding is continued, another critical point is reached where flow itself reaches a maximum and then falls off. The ultimate position is reached when there is a complete traffic jam with zero flow; cars are crowded as near to each other as possible, but there is no movement.

There is no clear measure of capacity to be seen in this. The point of maximum flow might be regarded as in some sense a norm, but there is no obvious economic justification for this. The optimum may well be at higher speeds than those possible at the point of maximum flow, though it cannot be at lower speeds. This can be shown as follows. We assume that people are ready to pay for time saved, and that this is the only variable cost involved in the use of the road.[1] At any given speed of traffic there is a normal demand curve representing the fact that people gain varying benefits from the right to travel on the road. The optimum point is clearly where the losses of those who are forced to reduce speed as the flow of traffic increases are equal to the sums the extra road users would pay rather than be forced off the road. We cannot make any general statement about what is the economic optimum, as we do not know the elasticity of demand for use of the road at any given speed, or the elasticity of demand for speed. It is however clear that the optimum cannot be at points where both speed and flow are below those possible at the point of maximum flow, because at this point the total number of effective users falls.

---

[1] Those who do not like a definition of marginal cost in terms of opportunities foregone can make the same point in terms of the optimal conditions of exchange. (See Little, *A Critique of Welfare Economics* (2nd ed.), pp. 129–33, 152–3; E. J. Mishan, 'A Survey of Welfare Economics 1939–1959 (*Economic Journal*, June 1960), pp. 206f.). There are cases where rationing is more appropriate, but this does not seem to be the case here.

To illustrate arithmetically: if, with 100 vehicles passing a point per period, the addition of another 10 vehicles passing in the same period, reduces speed by 1 m.p.h., the question is whether the extra ten vehicles are ready to pay, not merely the cost of travelling at the reduced speed, but also the extra cost incurred by the reduction of speed of 100 vehicles by 1 m.p.h., which is the value to those users of the time lost. This is the marginal cost of their use of the road.

There is no reason to assume that all road users want to save time, though it may be plausible to assume that most do. But there is clearly no reason to assume that all value time in the same way, or that their valuation of time saved is related in any simple way to their valuation of the right to use the road. For example, those who set a high premium on time-saving as the chief consideration, will clearly consider alternative methods of travel, such as train or air. Even allowing some convenience-value to use of the road, the possibility of alternative methods of transport at given speeds will set limits to the amount they are ready to pay for use of the roads. At the other extreme, the road user who uses the roads for convenience, and is not bothered by time, for example the family which goes out into the country at the week-end, and has no particular objective, may be ready to pay highly for the right to add to traffic congestion, but little for the advantages of speed. There are also users who are ready to pay highly for both. There may even be road users who would pay positively for slowing down traffic. The valuable studies of the Road Research Laboratory on time-saving and its costs only tell us a very little about these things.

The matter is further complicated by the fact that traffic does not flow evenly at a given speed, but consists of a varied assortment of people wanting to travel at different speeds. The actual flow possible will depend on the composition of this assortment.

---

[1] This, of course, is not the case. Car operating costs rise with delays (though they may be higher at higher (even) speeds). These costs can be calculated (see W. H. Glanville and R. J. Smeed,' The Basic Requirements for the Roads of Great Britain ' (Institution of Civil Engineers, 1958), paras. 35–56).

If we admit that we do not know anything about the elasticities of demand for speed and other variables, and that nothing can in practice be done to make the marginal road user bear the costs he creates for others, we can still come nearer to a sensible solution by arbitrarily setting either a minimum speed limit, or a maximum flow of traffic to be achieved by limiting the number of vehicles trying to use a road. There may be many cases in practice (ignoring peak problems) where the actual flow is reduced below the maximum by too great a number of vehicles trying to use a road; in these cases a limitation of use would obviously increase economic welfare. It may also be argued that in other cases a limit below maximum capacity may be justified on commonsense grounds. Thus it could be argued that users of trunk roads mostly use them because they want to save time, and that those who do not want to save time can find alternative routes which meander through the countryside. One might therefore be justified in limiting traffic on certain routes so as to achieve certain speeds.

Once a decision of this kind is made it should be possible to achieve the required aim by setting a suitable toll. Setting a minimum speed-limit will not by itself be practicable if there is no other way of limiting traffic and there is a tendency to over-crowd the road, because the speed will fall without anyone being able to do anything about it. Traffic will therefore have to be limited by direct physical coercion, or by toll, the second being obviously the most sensible. One advantage of a toll is that experience of its use will give us information about the actual demand for the use of particular road.

Once we abandon the pure case of an even and continuous flow of traffic through time, it is no longer possible to state in quite the same way that speeds and flows below the point of maximum flow are necessarily worse positions than the position of maximum flow. There will be fewer effective users, at the period (and place) of peak congestion, than there could be if maximum flow and higher speeds were achieved. But it is only if these conditions could be brought about without preventing people from continuing their journey outside the peak period and place that we can state that maximum flow or higher speeds are better positions. Thus, if, by holding up a certain number of people at a given point at a peak jam, we can increase both the number and speed of those allowed to move, this would clearly be an improvement, if those who will have to wait perhaps longer than they otherwise would have had to were selected according to their valuation of time. If, on the other hand, our procedures for maximising flow at the peak period and place actually reduce the number of users at other places and times, then it is no longer obvious that speeds and flows below those obtainable at the point of maximum flow are necessarily worse positions.

If we knew what people would be prepared to pay for the opportunity to travel at higher speeds, as well as the other operational costs of delays, we would know when to extend the existing road network, either by adding a new lane or building a new road. Ideally, we would compare the cost of these improvements with what those excluded from the existing congested road would be ready to pay in order to travel, plus what people would be ready to pay to travel faster than they can on the present road. A congested road may be making a surplus over its costs of construction, etc. We would continue improving roads until such a surplus disappeared. There may be no such surplus, where the cost of land is very high, even with a very high charge to limit traffic.

4. Under-Utilised Roads.

Many roads are clearly not used to their full capacity, whatever definition we give to full capacity. Quite different principles of charging may arise here. Users must clearly pay at least the marginal cost of their use of the asset. This includes the wear-and-tear involved in use, and also the maintenance (plus signalling costs, etc.) required if the road is to be kept open for users at all. The latter element need not be paid for by users according to their use, so long as it is somehow recovered from all the users collectively.[1]

Need there be any further charge made for assets of this sort? Under-utilisation in the sense in question is not at all incompatible with an asset making a profit and requiring replacement. In this case, the replacement cost of capital should properly be charged to present users; and the argument implies that this can be collected. In the case of under-utilised roads, it is perhaps difficult to see what would be involved in replacement that is not already covered by maintenance, etc. It might, however, be the case that some maintenance costs cannot be evenly spaced over time, but are incurred in lumps, say every ten years, and that this might be regarded as replacement. It may be that there are also minor improvements that would be justified but which can hardly be regarded as maintenance. All these costs should certainly be collected from users somehow or other.

There is, in this, however, no element of charge for the historical cost of building the road in question, or for the value of the land occupied by the road. How far should these elements also be covered? Historical cost is irrelevant. Replacement cost will allow for all capital elements that have to be covered. If there is opportunity to earn further profits to allow for historical cost elements that do not require replacement (everlasting assets, such as the cost to the Romans and Anglo-Saxons of diverting land to roads), there is no objection to their being earned, so long as they do not interfere with maximum utilisation. But if attempts are made to collect such profits by means of a charge related to use, they are likely to do precisely this, and should therefore be resisted.

Has the land any other relevant value which should be charged to users? Certainly, if it has a value in some alternative use, this sum should be collected from users, if the piece of land is to be maintained in this particular use. If one user cannot afford to pay as much rent as an alternative user, it is clear evidence that the alternative user can make better use of the land. If, therefore, local roads have a market value, for example on sale to neighbouring farmers, this is a real cost to the community which should be borne by road users. But it is difficult to believe that, in most cases of under-utilised roads, there can be any such value. (The cases of roads in towns, and where roads have access-value to particular property-owners, are dealt with later). In most, if not all, cases the value to farmers would be less than the cost of breaking up the road. Abandoned roads, like abandoned railways, are likely to have little value, and therefore there is no argument for charging any element of historical cost as a necessary part of the charge for their use.

---

[1] It must be remembered that external economies are ruled out, and that we are dealing with roads that are independent of other roads.

In the case of an under-utilised road, it is therefore desirable to ensure that no more than maintenance costs (including whatever replacement may be involved) are collected from users if extra charges are likely to affect use seriously. Though it might be fair to assume that the general elasticity of demand for use of roads is fairly low, it is possible that it is higher in remote country areas where roads tend to be under-utilised, partly because incomes may be lower there, and partly because transport costs are bound to bear more heavily on all aspects in life in these areas. Whether this is so or not, it is important to see if the running charges that the users have to bear in these cases are likely to be above or below something approximating to marginal cost of road use.

Some answer to this question can be obtained from the Road Research Laboratory censuses of traffic-flows. It is possible to compare the annual flow (in terms of vehicle-miles) per mile of road of certain kinds with the maintenance cost per mile and thus estimate the maintenance cost per vehicle-mile. This can in turn be compared with the incidence of the petrol tax. These figures are given in Tables I–III, in terms of very broad averages. The first part of Table I shows the distribution of traffic on roads of different kinds, no information being available for unclassified roads, which admittedly represent an additional mileage as great as the rest, 96,680 miles at March 31st, 1958 (as compared with the total of 94,470, when unclassified roads are excluded) of which roughly a third would be in urban areas. The table very clearly reveals the well-known fact that a small part of the road system carries a disproportionate part of the traffic. Trunk and Class I roads were 30 per cent of the roads in question, but carried 76 per cent of the traffic, and Class III roads were 52 per cent of the total, but carried only 14 per cent of the traffic. If we take the extremes, we find that Trunk and Class I roads in urban areas represented 9 per cent of the roads, but carried 36 per cent of the traffic, while, at the other extreme, Class III roads in rural areas represented 46 per cent of the roads, but carried only 9 per cent of total traffic. (Car traffic does not show much variation from the position of total traffic.)

The second part of Table I shows the density of traffic on roads of different types, and illustrates the same point in another way, also giving the distribution of traffic between passenger and goods vehicles, and showing the

TABLE I

British Roads. 1958

Mileage and Traffic	*Urban*	*Rural*	*All*
Trunk (miles)	1,730	6,580	8,310
% of total miles	2	7	9
% of total traffic	7	14	21
% of car traffic	7	15	21
Class I (miles)	6,310	13,390	19,710
% of total miles	7	14	21
% of total traffic	29	26	55
% of car traffic	27	28	55
Class II (miles)	3,930	13,670	17,600
% of total miles	4	14	19
% of total traffic	7	4	11
% of car traffic	6	4	10
Class III (miles)	5,570	43,280	48,850
% of total miles	6	46	52
% of total traffic	5	9	14
% of car traffic	6	8	13
All (miles)	17,540	76,930	94,470
% of total miles	19	81	100
% of total traffic	48	52	100
% of car traffic	45	55	100
Vehicle-Miles per Mile of Road ('000s)			
Trunk			
All traffic	1,910	1,060	1,240
Passenger	1,350	800	910
Goods	550	260	320
Cars	1,170	710	810
Class I			
All traffic	2,240	940	1,360
Passenger	1,680	750	1,050
Goods	560	190	310
Cars	1,350	650	880
Class II			
All traffic	880	130	300
Passenger	630	100	220
Goods	250	30	80
Cars	470	90	170
Class III			
All traffic	460	100	140
Passenger	380	70	100
Goods	80	30	40
Cars	320	60	90
All			
All traffic	1,340	330	520
Passenger	1,000	260	390
Goods	340	80	130
Cars	810	220	330

Sources: Roads in England and Wales 1958–9 (H.C. 1959–60, no. 42). *Industry and Employment in Scotland and Scottish Roads Report, 1958.* (Cmnd. 706.)

Notes: The table excludes unclassified roads. Trunk roads are those entirely paid for by the central government. Class I roads are those subject to 75 per cent grant to local authorities, Class II 60 per cent grant, and Class III 50 per cent grant.

Vehicles include motor-vehicles, but exclude pedal cycles. Figures of traffic flow are taken from Appendix XVIII, Table X of *Roads etc. 1958–9*, and represent the results of a continuous census taken at 50 points. Mileage of roads as at 15.5.58 in Scotland and 31.3.58 in England and Wales, adjusted to an urban/rural classification.

TABLE II

Maintenance Costs per Mile of Road (£) (1957–8)

	London	County Boroughs	Counties	All
Trunk				
England	—	2,570	1,750	1,540
Wales	—	2,260	1,290	1,130
Scotland	—	1,580	740	740
Class I				
England	4,860	1,810	990	1,170
Wales	—	1,700	810	860
Scotland	—	3,120	410	600
Class II				
England	1,820	1,170	580	640
Wales	—	1,610	490	520
Scotland	—	1,220	190	320
Class III				
England	1,560	1,010	310	340
Wales	—	810	250	250
Scotland	—	1,430	200	250
Maintenance Costs per Vehicle-Mile (d.)				
Trunk				
England		.3	.4	.3
Wales		.3	.3	.2
Scotland		.2	.2	.1
Class I				
England		.2	.3	.2
Wales		.2	.2	.2
Scotland		.3	.1	.1
London		—	—	.5
Class II				
England		.3	1.1	.5
Wales		.4	.9	.4
Scotland		.3	.4	.3
London		—	—	.5
Class III				
England		.5	.8	.6
Wales		.4	.6	.4
Scotland		.8	.5	.4
London		—	—	.8

Sources. Roads in England and Wales 1958–9 (H.C. 1959–60, no. 42). *Industry and Employment in Scotland and Scottish Roads Report, 1958.* Cmnd. 706.)

Notes: Maintenance costs for year-ending 31.3.58 for England and Wales, and 15.5.58 for Scotland. Maintenance includes repair and minor improvement, the latter including resurfacing, patching, new traffic signals, etc., within existing highway boundaries. It excludes cleansing, watering and snow clearing.

predominance of the private car. Very roughly passenger traffic represents 80 per cent and the private car 60 per cent of the total. The variation in flow per mile of road is between 100,000 vehicle-miles per year for Class III rural roads and 2,240,000 vehicle-miles per year for Class I urban roads. The division of categories of road is very broad and the sampling errors for particular categories may be high. In reality of course the variation of flow is very much greater.

The first part of Table II gives the variation of expenditure on mainten-

TABLE III

Petrol duty cost per mile for cars of different size (d)

C.C.	
Up to 1,000	.78
1,001 to 1.265	.97
1,266 to 1,500	1.12
1,501 to 1,750	1.20
1,751 to 2,000	1.11
2,001 to 2,250	1.21
2,251 to 2,500	1.29
2,501 to 3,000	1.40
Over 3,000	1.75

Source: Automobile Association, Nov. 1958.
Notes: Petrol at 4s. 10d. per gallon on average.

Relationship between traffic flow and road maintenance cost

Cost per Mile of road	*Costs per vehicle-mile (d) with given traffic flows in vehicle-miles per year*			
	.5d.	*1d.*	*1.5d.*	*2.0d.*
£100	48,000	24,000	18,000	12,000
£150	72,000	36,000	27,000	18,000
£200	96,000	48,000	36,000	24,000
£250	120,000	60,000	45,000	30,000
£300	144,000	72,000	54,000	36,000

ance of roads of different kinds by local authorities of various classes. The distinction between county boroughs and counties is not the same as that between urban and rural, and the national figures of annual vehicle flow do not give information for London, where flows are much higher than in urban areas generally. Maintenance costs include repair and minor improvement, but exclude cleansing, watering, snow clearing and administration. (The bulk of the latter items, with the exception of administration, was in fact spent by county boroughs, so that it does not seriously affect our arguments about under-utilised rural roads.) Maintenance costs therefore include rather more than the necessary minimum expenditure required if roads are to be kept open, and a great deal more than marginal use-cost. The figures reveal that maintenance costs per mile of road vary very widely, and that the variation roughly corresponds to the variation in traffic-flows, for broad categories of roads. (It may be that this merely reflects the different widths of these roads).

The second part of Table II gives the figures we are interested in. They do not represent the real maintenance costs per vehicle mile in the different regions, but what the costs would be if the flows of traffic in these regions were the same as the national average, and if expenditure per mile on urban roads reached the actual figures of spending by county boroughs (large burghs for Scotland) or London, and if that on rural roads was the same as spending by counties. There is no reason to expect that these conditions are exactly fulfilled; traffic flow in reality is much higher in London and lower in Wales and Scotland. The fact that the London figures are higher than the rest does not therefore tell us anything significant about the real costs. What is perhaps surprising is the degree to which variation in maintenance costs is evened out by the inverse variation in traffic flows. There appears, however, to be a significant difference between the expenditure on rural roads of Class II and those of Class III, which is not at all parallelled by the difference in traffic flows. Expenditure is roughly twice as high per mile of Class II road in England and Wales (but not Scotland) whereas vehicle flows are only some 30 per cent higher. As a result the highest vehicle-mile cost is to be found on these roads in England and Wales.

We are interested in Class III rural roads as the nearest approximation to under-utilised roads. With expenditure on maintenance varying from £200 to £310 per mile of road, costs per vehicle-mile at national average levels of traffic flow would vary between .5d. and .8d. Excluding the Class II rural roads in England and Wales, these are the highest figures to be found anywhere in the Table. They can now be compared with the figures of the cost of the petrol tax per mile in 1958, as given in Table III, and it is clearly seen that this is much greater than these maintenance costs, which are, as we have seen, themselves more than any true marginal cost.

To express the matter another way, the incidence of petrol tax at 1d. per mile would just cover maintenance costs at £200 per mile of road if traffic flow were as low as 48,000 vehicle-miles per year, or half the national average for Class III rural roads. The second part of Table III gives some examples of the relationship between maintenance cost per vehicle-mile and the two variables of traffic-flow and cost per mile of road.

Some idea of the position on unclassified roads can also be given. Maintenance cost per mile in 1957–8 was £130 for Welsh counties, and £160 for English counties (average costs for all areas being £200 for Wales and £260 for England). It would be useful to know whether there are many roads where annual vehicle-miles are as low as 30,000, and whether maintenance costs deviate much from the average, i.e. whether costs often rise above 1d. per vehicle-mile.[1]

It seems fairly clear from these figures that there may be a large mileage of roads in the country which are under-utilised, and yet where the petrol tax more than covers all the proper costs which should be borne by road users, who will at the same time be paying rates, licence duties and purchase tax. Rural dwellers would seem to have some ground for complaint about these facts.

---

[1] A survey of traffic in four areas in 1954 (*Road Research 1956*, Department of Scientific and Industrial Research, p. 23) shows a weekday flow of 200 vehicles including pedal cycles, of which 122 were motor vehicles. Grossing up by 300 days gives an annual flow of 36,600 vehicle-miles. Since 1954 traffic has much increased, and this makes no allowance for Sunday traffic.

The chart on the following page shows that only 30 per cent of roads in Dorset had daily flows of 100 vehicles or less (30,000 per year), 18 per cent in Leicester C.B., 10 per cent in Lancashire, and 7 per cent in Bucks. 10 per cent of roads in Dorset had flows at half this level and 5 per cent in Leicester C.B.

1956 data (*The Basic Requirements for the Roads of Great Britain* by W. H. Glanville and R. J. Smeed (Institution of Civil Engineers, 1958), p. 20, Table 3) show weekday flows on rural unclassified roads of 140 motor vehicles, making annual flows on our basis of 42,000.

This information is strictly of flows past scientifically-selected points (see J. C. Tanner, 'The Sampling of Road Traffic', *Applied Statistics*, Vol. VI, No. 3, 1957, pp. 161–70).

We have, however, gone far beyond what is justifiable on our principles, which exclude any interdependence of the road system. If, in fact, the road system is interdependent in the sense that the users of under-utilised country roads are also the creators of traffic congestion on main roads, then it is no longer inappropriate to charge them more than the maintenance costs of the rural roads. High petrol tax on rural roads may be reducing congestion on other roads. In practice, it is important to know how really interdependent the road system is, so that we could isolate parts of the system which are relatively independent. Without such knowledge, generalisations about under-utilisation can only be very rough.

5. Access Roads.

Let us take the case of a pure access road to a house or group of houses, which gives on to a trunk road.[1] The use of space for the road is complementary to the use of space for the houses in question. The buyer of space for a house wants not only a suitable site for house, garden and garage-space, but also access, so that he can travel from his house other than by foot, and so that others can visit him or deliver goods to his house. The road is assumed to be of no use except for these purposes. The developer of the house will naturally build an access-road when he builds the house and the occupant will only want to buy the house if there is access. The value of the house cannot be estimated without the road, no more than can that of the road without the house, though they can be valued jointly. The demand for house and road is complementary, or, to put it in other words, the use of house and road is a joint demand. There is a value for each when the cost of the other is given.

If we suppose that the access road is provided free, clearly the whole value of the joint product accrues to the house-site. In so far as costs are involved in the provision of access roads (and these costs include both the original extra land requirement and the costs of road construction and maintenance), free provision of roads leads to a higher value of the house-site than would otherwise be the case. A developer would be able to develop at a lower density and more space would be available round houses for the same price.

Let us now suppose that the cost of roads is paid for out of an equiproportionate tax on the value of all land and buildings. It may be that the cost of providing access is not equal in proportion to the total value, perhaps because it requires more space in a particular case, or because topography makes road construction and maintenance more costly. If access is provided free, there will then be a greater rise in the value of the house-site over what the value would otherwise be than in other cases. Then clearly those whose roads cost more than the average are being subsidised by those whose access roads cost less than average.

---

[1] The argument here assumes that each site has a piece of road allocated to it, but this is of course unrealistic. Access, parking and traffic flow are both substitutes and complementary, and normally shared between several sites.

TABLE IV

	Case A	*Case B*	*Case C*
Value of developed site	100	100	300
Cost of Road	10	20	10
' Real value ' of site (Value less cost of road)	90	80	290
(1) Road-rate levied on A and B	15	15	—
Net value of site	85	85	—
(2) Road-rate levied on A and C	5	—	15
Net value of site	95	—	285

Similarly, those cases where total value (of developed site) is higher in relation to the cost of access roads subsidize those where it is lower, if a proportional tax is levied. These two cases are illustrated by Table IV.

The value of the access road can be given in reverse as the value of the site less the cost of the building on the rest of the site. It may be that it is also meaningful, in some cases at least, to assess a value of the site minus access rights. (Some people may be happy to walk to their houses and pay to have deliveries, etc., carried by hand). Whichever way we look at it, a monopoly road provider could extract some access-value from users, and there is no reason to suppose that this access-value is equal for all users even where total value is the same and the cost of providing roads the same. If, in such cases, the monopoly provider charges the same to the two users who value access unequally, we could say that this is cross-subsidisation of those who value it more by those who value it less.

It is not merely that some people value access rights more highly than others, but that access rights are more valuable in some places than in others, the places where certain particular land-users settle to whom access rights are all-important. Thus this ' cross-subsidisation ' leads to a greater inflation of site-values at certain points than at others. This is clearly of more importance to commercial users than to others. We have seen that an equi-proportional tax (which is in effect the way the rate proportion of road cost is levied) makes the high value user pay more than the cost of providing the road, if this cost is no higher than elsewhere. His contribution will be more than a tax based on the monopoly value that could be extracted, if the ratio between access values is less than the ratio between total values, and vice-versa. Let us take the above cases A and C, valued at 100 and 300, with road tax at 5 and 15. If the access-values respectively are, for example, 50 and 150, then a monopoly tax affects both in the same proportion, and, scaled down to the level of road costs, would be the same as the proportional tax. But if the access values are 50 and 100, then the high value property still pays more than its share, if this monopoly criterion is regarded as setting any kind of standard.

Among the proper costs of the road are those incurred by other road users as a result of congestion. Thus, if a road is used both for access and for traffic, as is the normal case in shopping streets in towns, congestion costs must be added to other road costs. It is clear that the commercial land user both creates the major part of the congestion cost (with his delivery vehicles, or customers) and has the highest site values. In terms of ordinary road cost, there may be some increase above average levels in the case of shopping streets, but this is unlikely to be as great as the difference in road-rates paid on site-values. But when congestion costs are taken into account, the picture may be very different. Will the ratio of costs still be less than the ratio of values?

Road-rates levied on property values only cover a part of the cost of roads, that part not paid for by central government. Unclassified roads are wholly paid for by local authorities, and there were 40,000 miles of them in Great Britain in 1958, out of a total 57,200 miles of road in all areas of urban authorities. Classified roads receive varying grants from central government, but even unclassified roads are subsidised from central funds, in so far as the block grant would be different were local authorities not responsible for their upkeep.

It is sometimes argued that cities must do something to improve their roads or provide parking and other facilities for motorists, presumably at less than cost, if they are not to avoid loss of rateable value and a decline of business. On the contrary, it is clear that site values which depend on access are higher than they would be if cities treated access-values as would a commercial landowner who charged what he could get for the use of his land. Even if we ignore this, and consider merely the costs incurred in road use, it is almost certain that the commercial user pays much less than the costs he creates. If he caused no more traffic congestion than the other frontagers, and if all road costs were levied by means of a road-rate, he would pay more than his proper costs. The fact that rates only pay for some of the out-of-pocket road costs means that other frontagers pay much less than their proper costs, but does not necessarily mean that the commercial user pays absolutely more than the out-of-pocket costs. It would be interesting to know more about all this.

6. PARKING.

A permanent allocation of land to parking should be charged at least the economic rent for the use of the land, that is to say the value someone else would be ready to pay rather than be deprived of the use of the land. This is the cost to the community of the diversion of land to the purposes of parking. The principle applies equally to land used for the building of garages, for land set aside for parking by means of parking meters, or even for land permitted to be used for this purpose by that most arbitrary of planning authorities, the local police. In all these cases, land is diverted from other uses, and should be charged for per square foot.

The charges for parking in parking meters are no different in principle from the charges that would be charged commercially by a garage. That is to say, the lowest charge should be that which will cover the rent which could be gained by letting off the land to another user, plus the cost of fitting up parking meters, and the cost of administration of the system. An upper limit would be set by the demand for the use of the space, and, if garages are provided commercially at the same time, competition will prevent parking meters earning monopoly profits. If, for some reason, this competition does not work adequately, presumably a public authority would prefer a system which made maximum use of the space available for parking rather than a system which maximised profits, if there were a conflict between these objectives, and would seek to expand space for parking, so long as revenue covered the costs and rent.

So far we have assumed that parking uses up land which would be available for other purposes for which a rent could be charged. But what if, in fact, the land could only be used for traffic or parking? The rent here in question is then the price the mover will pay to have traffic speeded up by the removal of the congestion caused by the parker. (It is of course possible that the parker will pay more to have the traffic congested, though it may be doubted).

The parker we have considered so far is the long-term parker, not the deliverer or visitor (shopper), who wants access rather than road space, in so far as we can make a distinction. We can ask the same question of him as we ask of the parker, and, as an alternative to charging the site-owner for the access-right granted to him, we could charge the visitors. A shopper will be ready to shop at a particular shop, in spite of a visiting charge, because prices are lower there; if the shop has to pay access-charge, the shopper will have to pay for it in the long run. Not all shoppers come by car, however, and it would be fairer to charge the shopper. But administratively, it may be much easier to collect an access-charge than a visiting charge.

Bus-stops are merely a particular form of stopping right, which could easily be treated as parking spaces reserved only for buses. They should also be charged for in the same way. Clearly, however, the incidence per person is much less than in the case of the parking charge for a car.

7. The Peak Problem and the Urban Commuter.

To demand the use of the same asset at different times is not necessarily to demand the same service, nor does it necessarily involve the same cost. Hitherto, we have treated roads as either congested or under-utilised, but this may apply to the same road at different times of day, most notably in the case of the urban commuter, but also in the case of the bank holiday week-ender. Rational use of resources will require a different price for use at different periods. Commuters who create congestion costs should pay for them, and they should pay for extra capacity of roads provided to meet their particular needs. If this extra capacity is provided to meet such a peak demand, and charged to other users or paid for out of taxation, there is no guarantee that there will not be a great waste of capital.

The problem of the urban commuter exemplifies most of the complexities of the road problem at their worst. It is the cheapening of urban transport that has caused cities to grow over a wider area, and thus benefit from the economies of scale that this enables firms and industries to reap. High site values reflect these economies of scale resulting from congestion of employment and shopping in central areas. In so far as the cheapness of transport facilities to the urban shopper or commuter does not reflect their real cost, these economies of scale are pressed beyond the point that is economic to the city, and site values are raised to excessive heights. But this process could not occur if people could not get to work in the central areas. Public transport facilities are therefore run at a loss, because they cannot compete with the car, and congestion grows, while parking becomes ever more of a nightmare. New roads are not built, because of the prohibitive cost.

The private car commuter pays, as we have seen, in the form of a slower movement than he might be ready to pay for. But he does not pay for the marginal cost of the congestion he causes to other commuters. As a parker, he does not pay rent for the land he uses on an economic basis, and, in so far as new roads or road improvements take place, he does not pay for them according to use. If the urban commuter paid the full economic charge for all these things, we would be in a better position to know whether to invest in new urban roads.

It may be that, in the case of some urban roads, there is a real economic rent that could be charged for the use of some existing roads. That is to say, there may be some urban roads that would have a real use to adjoining site owners. In so far as this is so, these roads should be charged a price that at least covers that rent. Otherwise, the value of the land used for roads is only relevant when a new road is in question. Whether road users would be ready to pay the cost of a new road would become clear as prices were raised to cover growing congestion costs. At some point the revenue will be sufficient to pay for the enormous costs of the new roads in urban centres.

Nor must it be forgotten that the peak may shift. The worst peaks in urban transport are due to the conventional hours of work of offices and shops. Efforts to reduce the high peak costs by staggering hours voluntarily have led to little results. It may be suggested that the root reason is not that the conventions are too strong, but that the economic forces are too strong. That this is a plausible analysis is confirmed by the report on '*Crush Hour*' *Travel in Central London* produced by the Committee for Staggering of Working Hours in Central London (1958), which found that 72 per cent of firms were not ready to stagger hours because of loss of business efficiency and 23 per cent because of staff difficulties.[1] At the same time the Committee rejected 'a proposal for the charging of higher fares during the peak hours', because of 'strong public resistance', 'many objections on the grounds of practicability', and 'bearing in mind that it would penalise those workers whose employers would not agree to work staggered hours and also those workers who would be obliged to remain in the peak periods to preserve the balance of traffic'.[2] But it is not clear why, if it is the higher cost that deters firms from staggering hours they would not be amenable to pressure from the cost side; nor is it clear that workers have no choice of jobs and would be unable to press for higher wages to compensate for inconvenient hours of work. Whether or not any such scheme is practicable, it is not obvious that demand for peak hour travel, based on conventional working and shopping hours, could not be severely reduced and largely staggered by charging properly for the costs of peak hour congestion.

---

[1] P. 18.

[2] P. 24. They also thought that 'shopping hours could hardly be other than uniform' (p. 29), ignoring the experience of other countries.

If, in fact, a fairly moderate raising of charges for peak use of roads led to considerable staggering, it would be quite absurd to build new urban roads at great cost to meet urban congestion problems. So long as peak road users do not pay their proper cost, there will be constant pressure for more roads, at great cost both in money and in dislocation to city life, without any real knowledge as to how intense is the demand for them. It would be reasonable to suppose that in fact a good deal of urban road improvement would pay for itself, but it is equally reasonable to suppose that much of the demand for peak-travel by private car and much of the rigidity of working and shopping hours are encouraged by the subsidisation of peak use of the roads.

8. Planning and Price Solutions.

It is not implied in what has been said already that solutions of road problems by charging for the use of the road are necessarily more practicable or desirable than a solution by planning. The advantage of considering the various problems in terms of the right charging policy is that it brings the fundamentals to the fore. It may, for example, seem obvious to some people that speeding up traffic is a desirable end in itself, and that a simple planning solution (e.g. a Pink Zone) will achieve this end. But a consideration of what people would be ready to pay for speed, parking and delivering may suggest that this is not necessarily the best solution, and that people would prefer on balance to pay more for travel at a lower speed. (If, however, a guess could be made at an answer to such a question, and if we were sure that people really preferred to travel, shall we say, at 25 m.p.h. in London, then we could produce this result by planning, and perhaps more easily.)[1]

It may be that planning solutions to many problems of urban congestion are better than most. It is certainly worth considering how many of these problems are due to failure to use the price mechanism at crucial points. If we believe that improper allocation of peak travel costs is a major factor in causing congestion and that the economic loss in staggering the hours of businesses would not be great, a planning solution by forced staggering of hours of work would become sensible, and would obviously be better than to build new roads. Furthermore, if a planning solution is considered to be right for this sort of reason, or because it involves matters that cannot be subjected to the price mechanism (such as amenity, social costs of various kinds etc.), it still remains possible to produce such an end by the use of prices, as for example by means of a tax. It would be perfectly possible to reduce urban congestion by taxing employment in given central areas, or to enforce staggering by taxing firms that refuse to stagger their hours. Such a method of enforcement of a planning solution might be more effective than an alternative. There is certainly no sense in having planning and prices pulling in opposite directions.

---

[1] Nothing is said in this paper about bicycles and pedestrians. Clearly similar questions can be asked of their use of roads. The case of pedestrians, however, may be very suitable for 'planning' solutions. The failure to guide (and discipline) pedestrians in Britain is one of our major follies.

In general it has been assumed, however, that there are no great questions of social cost involved in the roads that cannot in principle be looked at in terms of ordinary money costs. This is not entirely true. The cost of road accidents to the community may be much greater than the cost in insurance to the car-owner, even if car insurance were profitable to insurance companies. Building of roads may involve loss of amenity in country areas; certainly turning the centre of cities into networks of fly-over crossings, with car-parks in the spaces in between, is not one that appeals to most people (though it may be doubted whether such a form of city is one that pays for itself economically). Nor is the noise and air-pollution brought about by the motor-car something that we should tolerate too easily. These issues are rather obvious, if very intractable. But it is sometimes argued that roads (and perhaps transport generally) produce such great external economies that they should not be treated as ordinary economic assets at all. It is this argument which, I would argue, we should examine most critically and closely.

9. External Economies.

The ordinary external economies argument is that an activity produces benefits to the community which do not accrue to the promoter of the activity, and which he cannot therefore in any way charge for. The clearest kind of case is that where the growth of a firm in a particular place leads to economies of scale in other firms providing some service for it (for example, making it economic for a lawyer or accountant to move into the town and provide a full-time service to the community), or to similar economies in the labour market. The firm, which makes this possible, benefits from the new low cost, but cannot normally in any way benefit from the gain to the other firms. This kind of case is perfectly general and applies to transport and roads as well as other activities. But it is to be noted that it is a question of reduction of costs in factor markets, inputs to the industries, not through the output of the firms in question; that is to say, in the case of roads, the argument would be that road-building, and the use of roads, would reduce the costs of the factors they employ. No doubt, in many cases, a new road leads to some reduction in the costs of civil engineering firms, from which other industries may benefit, and its use leads to some spreading of costs of the police force where this was formerly under-occupied. But it is not such cases that people have in mind when they associate transport with great external economies.

The real argument is that the building and use of roads produces benefits to the community which cannot be recouped (in principle) from those who use it, whether travellers or road hauliers. All possible benefits must accrue through a reduction of transport costs; it is assumed that roads are not regarded as objects of value in themselves. The question then becomes whether users are, for some reason, unable to recoup from the ultimate beneficiaries. It is not denied that it may be impracticable to charge users according to the benefits they gain; our argument is concerned with what such charges would be like, if it were possible. But it is difficult to see that there are difficulties of anything like the same magnitude in the case of the immediate road users charging the ultimate beneficiaries. If the opening of a new road enables farmers to develop new crops near a market, it is difficult to see that this will not be reflected in the farmers' demand for transport.

It is of course the case that producers of an intermediate good or service will not be able to recoup all the benefits that accrue to their customers. They may be unable to practise discrimination on an adequate scale, or they may be inhibited by institutional arrangements. If we grant that this may be important in the case of roads, we still have to meet the argument that this applies in general to all intermediate goods and services which are used by widely different customers. Is there any reason to suppose that the benefit to the community from a new or improved means of transport is greater than the benefit from an improved supply of electricity, wood, steel or paper? Is there any easy way in which it can be shown that the producers of these commodities can recoup external benefits in ways which are not open to providers of transport? If this cannot be shown, the external economies argument for charging below cost in the case of roads becomes a general argument for subsidising all intermediate goods and services which are widely used in the economy.

The building of a road involves great indivisible expenditure, and the return on the investment may take a long time to mature. But it is not clear that these facts, which are often in the minds of those who talk of external economies, are strictly relevant. They suggest that building a road is a matter for very wide consideration, and rather inevitably a matter for governments and not private enterprise. But because a lot of money is involved, and because the return may only mature when a host of new industrial developments have occurred in the course of time, there is no argument for not charging for the use of the road (perhaps not in the early days), or for not trying to cover cost (in some perhaps long period of time). The risks may be very great, the estimates very faulty. But the matter is not in principle different from other forms of investment. It provides no case for a subsidy, and does not lead to any suggestion that a road that the users could not be made to pay for (if an ideal charging system could be devised) would benefit the community more than a power station, which they would be ready to pay for.

The crucial question is whether there are any benefits accruing to the community from the use of roads which cannot in principle be charged for, and, if so, what these benefits are. I must admit that I cannot see that there exist any benefits in this field different from those that occur in the case of any similar industry producing intermediate goods and services for a wide range of other industries.

I suspect that people are obsessed by the influence of transport on the location of industry, perhaps quite rightly. I assume with contemporary opinion that the market does not produce a pattern of location of industry, which maximises national output, and that this failure shows itself in unemployment in certain areas and a discrepancy between the costs of the private firm and the costs to the community such that, while production costs of a firm are lower in one place, this may be offset by higher costs of houses, public utilities, social services, etc., which are not reflected in the firm's budget. Sometimes these costs may be money costs, or at least calculable in principle in money terms; in other cases they may be pure social costs. I assume therefore that it is right to subsidise production in some parts of the country and check it in others, in order to offset these discrepancies. This being so, there is no reason at all why transport should be treated differently from other forms of economic activity; transport is as much a proper claimant for subsidy as any other industry, but it has no better claim.

The problem of a subsidy for encouraging the proper distribution of industry is one of steering development in the right direction with the least expenditure of public money; to maximise output at lowest cost still remains the aim. It is in fact reasonable to suppose that, of all the influences on patterns of industrial location, transport is one of the most important, even if it may be difficult to estimate its effects precisely, because they are so widespread. But if this is so, it implies that one of the cheapest ways of distributing industry in the right pattern may be by providing transport facilities at less than cost, or at least taking great risks in the provision of facilities ahead of demand.

In this latter sense, which does not differentiate transport facilities from other forms of production, there may be a case for subsidising transport, but there do not seem to be any other situations where it is relevant. It may well be that our transport planning authorities have been far too unimaginative in this respect, and have thereby frustrated the actions of the Board of Trade and the planning ministries in trying to decentralise and disperse industry. The failure to build suitable roads and bridges may have been particularly hampering here.

10. Methods of Charging.

If we want to apply a pricing system to the roads, we are not entirely without means to do so. In fact we have a much greater variety of methods of charging than would be possible in many industries. It is perhaps worth outlining them, with some comment on their usefulness for particular purposes.

(*a*) *Vehicle Taxes.* Clearly these are most useful if an assessment can be made of the depreciation costs particular vehicles cause by their movements on the roads. If congestion-causing characteristics can be attributed to particular sorts of vehicles, as has been done, this again can suitably be charged to the vehicle. But in general the use of vehicles of a given type shows great flexibility and variability, so that vehicle taxes cannot be closely related to road use or road costs.

(*b*) *Taxes on Owners of Vehicles.* If overhead costs are to be levied in a discriminatory way on owners of vehicles, so as to prevent under-utilisation of roads, it would be possible, at least in the case of private cars, to tax the owner according to some other standard than the value of the car (as is done with purchase tax). Car owners could be taxed on their income-tax bracket or the rateable value of their house (compare the standing charge in an electricity tariff). If the government were worried about collecting a tax on ' luxuries ' (and no doubt popular feeling about motoring taxes looks at them in this light), this might be the most sensible way to collect them.

(*c*) *Taxes on Property-Owners.* This is easily collected and is relevant to the question of access-value. But these taxes are not collected today in any very sensible way in relation to road costs.

(*d*) *Tolls.* Experience in America suggests that tolls are easy to collect, readily acceptable, and a sensible way of paying for major roads and bridges. If they could be varied, they could be used, not merely as a means of assessing elasticity of demand for the use of particular roads, but also as a means of dealing with peak flows.

(*e*) *The Petrol Tax.* Much of the above discussion implies that the petrol tax, from which most money is collected from the motorist, is perhaps the least suitable of taxes, unless it can be clearly shown that the elasticity of demand for further travel is fairly low, so that a high petrol tax does not inhibit the use of roads where they are under-utilised. It is particularly damaging in its effects on business travel and road haulage, where the elasticity of demand must be fairly high.

If we want a charging system that does justice to the variety of demands made on the roads, and the varying pattern of utilisation, we should be ready to vary the taxes levied. Would it not be possible for example to have different vehicle licence duties for towns and rural areas? Would it not be possible to have special peak charges for use of certain streets, e.g. in the Pink Zone in London? Why should parking charges be fixed once for all, or be the same for all parts of one city? There are many possibilities.

11. CONCLUSIONS.

Three major questions for discussion arise from the above:

1. How far does it provide the right sort of framework for analysing roads as economic assets?

2. What kind of information that is not already available would be

required to answer some of the questions on which policy hinges? How far is it possible to collect this information? What second-best information can be collected, if what we would most like to know is beyond our reach?

3. Does such a framework lead to any sort of practical charging policy? Does it suggest what would be second-best policies, if we cannot follow what might be ideal?

Nuffield College,
Oxford.

COVERING COSTS BY RECEIPTS

By M. E. PAUL

This comment is confined to Mr. Munby's pricing rules for new roads. He argues that the entire cost of such roads must be borne by those who use them. The ideal system, he says, would be one which, by discrimination or a two-part tariff, covers the indivisible costs out of consumers' surplus. Thus for the marginal journey no more than marginal cost would be charged, while receipts would nevertheless cover total costs. If, however, discrimination is impracticable, we may according to his rules be thrown back on average cost pricing, for covered in some way or other by users' receipts all costs must be.[1] Mr. Munby does not bring arguments to support these rules, but refers for their justification to an article by Mr. Lewis.[2]

We here try to establish the following points:

I. The covering of costs (preferably by discrimination) is not, in the case of new roads, a good general rule for pricing.

(*a*) It provides an investment criterion which can be justified neither by commonsense nor by economic theory.

(*b*) As a guide to output, it may unduly restrict the use of a road.

(*c*) Discrimination may be an undesirable way of covering overheads, since it is unfair to certain consumers.

II. Mr. Lewis' arguments, on which Mr. Munby's rules rest, are fallacious.

Since Mr. Munby confines the system of charging which we are here discussing to *new* roads, it seems reasonable to assume that he intends it to help decide when and where new roads should be built. Though he does not separate its two functions, the rule is to provide a guide both to investment and to output. Let us first consider whether it really would be a good general guide to investment in roads. Should new roads be built when and only when the entire cost of each road will be covered by receipts from its particular users?

There are commonsense reasons against this. Many new roads bring substantial gains and losses to people other than users. By-passes ease congestion on other roads, and may improve the living of large numbers who never themselves go on the new road. Any road through a built-up area is liable to give rise to questions of general amenity quite apart from its value to those who use it. Deciding to build a new road has often much in common with deciding how best to lay out a new town. Innumerable people may be affected, some users of the road and some not. It may not be possible to consider the matter solely in terms of how much people in general would pay for this and that arrangement, still less to consider it simply in terms of how much users would pay. Is it really thought that the building of a Meadow Road in Oxford should be settled by estimating whether those actually using it would be prepared to pay the cost? Are the gains of High Street users irrelevant? If we take Mr. Munby's rule seriously, this is what it seems to imply.

---

[1] Munby, 'The Roads as Economic Assets', p. 275. 'They (i.e. all the fixed costs of a new road) have to be collected somehow or other from the users. . . .'

[2] W. A. Lewis. *Overhead Costs*, Ch. 1.

Perhaps, however, economic arguments support it. Mr. Lewis, for example, seems to think that covering indivisible costs by receipts is in some sense necessary for the optimum allocation of resources. The following consideration however, counts against such a view. Optimum allocation of resources is achieved when precisely the correct output of each good is produced.[1] Chairs are to be made in such numbers that neither by producing more nor by producing fewer could one person be made better off without harming another. So are tables and road journeys. But the charging rule 'Cover costs by receipts', is compatible with innumerable alternative outputs, depending on the type of charging used. How many road journeys will be made if this rule governs charges will depend upon *how* receipts are made to cover costs. Average cost pricing will lead to one degree of road use and discriminatory charging to another. Depending upon how discrimination is carried out, use may vary all the way from what it would be with average cost pricing at one extreme, to what it would be with marginal cost pricing at the other. It can hardly be shown that with each one of these alternative outputs it would be impossible by producing a little more or a little less, to make one consumer better off without harming another.

The fact is that no rules can be given which would enable us to decide whether the allocation of indivisible productive units is or is not ' optimum '. Marginal analysis cannot be applied to indivisibles, nor can the problem be solved in terms of surplus criteria. Let us briefly illustrate why marginal analysis cannot be used here.

The rule that price should equal marginal cost, is, given certain assumptions, a necessary condition for optimum allocation. Only if this rule is kept, is it possible for the output of every good to be such that by altering it no one could be made better off without another being harmed. The rule is needed because by keeping it, the transformation ratio in production between any two goods can be made equal to consumers' marginal rate of substitution between the two. Unless these two ratios are equal, someone *could* be made better off without anyone else being harmed.

Suppose that given resources are allocated to the production of X and Y, and T T′ (Diagram I) represents the alternative combinations obtainable. I_1 and I_2 are consumers' indifference curves. It is plain that unless X and Y are produced at a point P where an indifference curve is tangential to T T′, this consumer could be made better off and no one else suffer if X and Y were produced in different proportions. If the slope of T T′ represents the marginal cost ratios of X and Y, by making price equal to marginal cost producers will produce at a point of tangency.[1] Hence making price equal to marginal cost is a necessary condition for an optimum.

---

[1] We ignore the possibility of alternative optima.

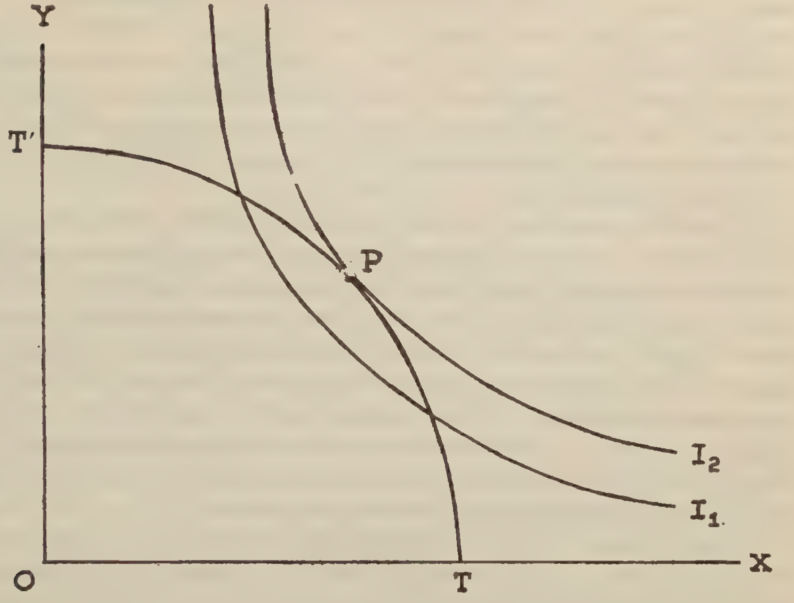


Diagram I

When, however, one of the goods can only be produced by an indivisible productive unit, this analysis cannot be applied.

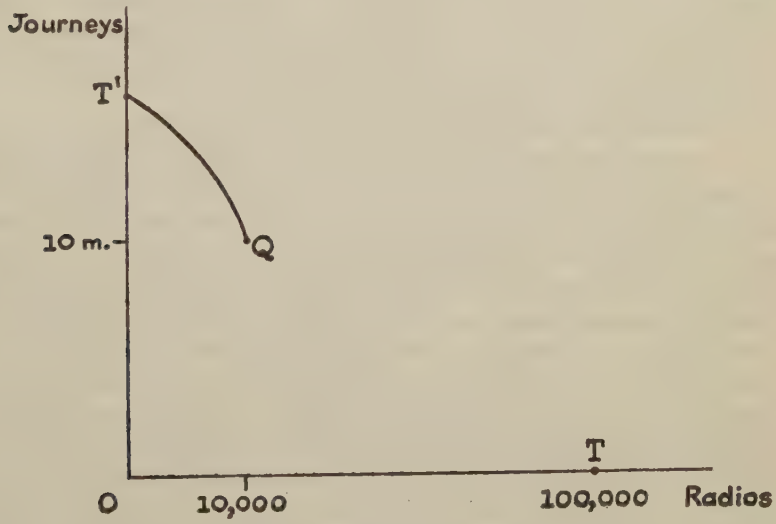


Diagram II

---

[1] We ignore here the difference between equality and proportionality.

Suppose we draw up a transformation curve for radios, and road journeys from A to B (Diagram II). It may be that the narrowest and cheapest road which could be built would enable 10 m. journeys to be made, and would require resources needed for 90,000 radios. So T T′ might begin at T with no journeys and 100,000 radios, and the next point be Q, with 10 m. journeys and 10,000 radios. Beyond Q a smooth curve to T′ might exist as the road is gradually widened. But between Q and T there is no curve, and hence nothing for an indifference curve to be tangential to. Thus marginal analysis cannot tell us whether the road should be built or not.

There is another way in which we can think of indivisibility in roads. That optimum allocation requires price equal to marginal cost can (in part) be argued on the following lines. Let us assume that price represents the value of a marginal unit to consumers, and that marginal cost represents the cost to society of that unit. Then by producing X and Y each just to the point where marginal cost equals price, we ensure that the value the consumer gets from the last unit of X is no greater and no less than the value he gets from the amount of Y which the resources producing the last unit of X could alternatively make. Each good is produced precisely to the point where the value of the last unit is worth its cost to the consumer.

When, however, we are deciding whether or not to build a road, we may be deciding between 1 m. journeys and none. The value placed by consumers on some of the journeys will be greater than that placed on others. The demand curve for journeys will slope down over the range from 0 to 1 m. So since we have to produce 1 m. journeys or none, it is not possible to produce journeys just to the point where the value of the last journey to the consumer equals its marginal cost. We have on the one hand an outlay so large that it may change factor prices, and on the other 1 m. journeys, some of which are of greater value than others. Thus where the smallest possible increase in output is so large that some units are worth more than others, marginal analysis cannot be used to show what output is optimum.

It has been suggested, however, that marginal analysis could provide a model for the optimum allocation of divisible factors, while surplus criteria are used to determine the optimum for indivisibles. On this system whether a road should be built is to be settled by surplus calculations, while how much use is made of it depends on demand at a price equal to marginal cost.

It is difficult either to define what is meant by the change in surplus resulting from an investment, or to say how it should be measured. Let us first define it as best we can, in order to illustrate the impossibility of measurement. The change in surplus resulting from investment in X is then taken to mean the sum of the compensating changes in income required to restore everyone affected to the level of well-being they enjoyed before the investment was made. If more X were produced, some people would gain, and others would lose. Some consumers of X, for example, would gain, and so would some consumers of goods complementary to X. On the other hand, some producers of substitutes, and consumers of goods which rose in price as a result of X expanding, would lose. For each gainer there is some sum of

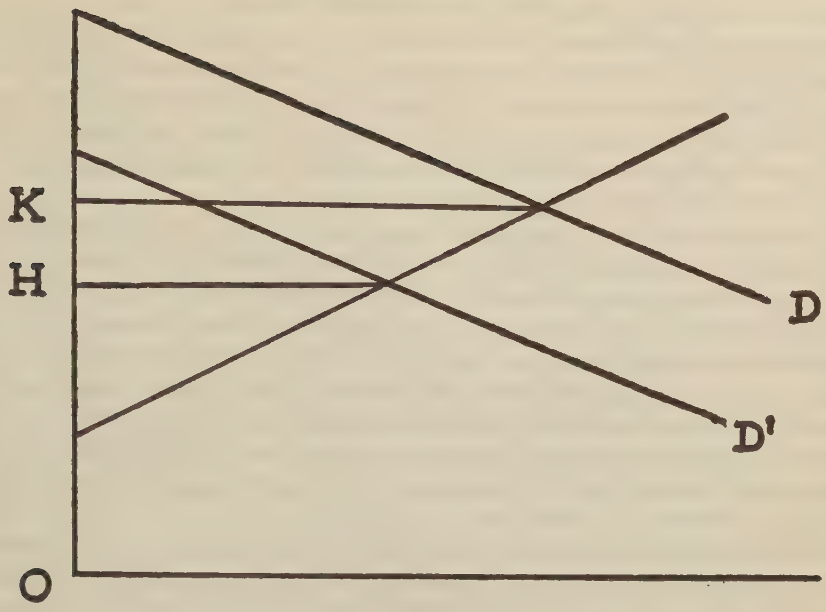


DIAGRAM III

money which we could take off him when more X is produced, and leave him no better and no worse off than he was before. And there is some sum which would compensate each loser for his loss. The net change in surplus due to investing in X is, then, the sum of these compensating changes in incomes for those affected; gainers' compensations less those of losers. This definition is itself unsatisfactory[1]; in particular it is open to the objection that the sum which could be taken off a gainer and leave him no worse off than before is not necessarily the same as the sum he would require to be bribed not to want the change.[2] However, it is the best we can do; we shall assume that it is sufficiently definite to be worth talking about.

It is plain that we cannot find out and then add together all the compensating variations in income for each individual when output of X expands; the hope of surplus analysts has been that these gains are sufficiently well represented by areas under demand and cost curves. This hope is not, however, justified. There are gains and losses which cannot be represented by areas under demand and cost curves.

Suppose, for example, we are wondering whether to build a bridge. We have in principle to measure not only surpluses arising from the bridge, but gains and losses of surplus on related goods. Let us consider the consumers' surplus from a tea garden, demand for which falls when the bridge is built. Let the price of teas fall from OK to OH if the bridge is built (Diagram III). Will there be any loss or gain of consumers' surplus on them? We might argue that since the area under the demand curve is now smaller, there will at least be no gain to consumers of teas to worry about. But this is not so; there may be a gain. For since the price has fallen, some may go and have teas who previously did not, and of these, some may enjoy a surplus. At OK they would not buy, but at OH they enjoy surplus. Their coming in stops the demand curve moving as far left as it otherwise would, but their surplus does not correspond to any part of the area under the demand curve.

---

[1] J. R. Hicks. The Four Consumers' Surpluses. *R.E.S.*, vol. XI, 1943.
[2] I. M. D. Little, *A Critique of Welfare Economics*, Ch. X.

Various attempts to solve this problem have been made. Professor Hicks has invented an ingenious step by step method of drawing up the curves for the bridge so that gains and losses on related goods themselves appear on the bridge diagram.[1] But for the gains and losses to be so represented, it is necessary that divisible factors should themselves be everywhere optimally allocated. If, as is the case, many conditions for an optimum, such as the absence of marginal taxation, or price everywhere equal to marginal cost, are not satisfied, then gains and losses cannot be represented by areas under demand and cost curves. This method does not give us a criterion for the optimum allocation of indivisibles.

We must therefore conclude that the theory of optimum allocation cannot at present be applied to indivisibles.[2] They are outside the scope of marginal analysis, and surplus criteria are impracticable. The rule that costs should be covered by receipts cannot then be justified as a necessary condition for the optimum allocation of resources. Roads are highly indivisible, and economic theory provides no rules for the optimum allocation of indivisibles.

Before we leave this question of an investment criterion for roads, there is one other line of thought worth investigating. It is possible to feel that since cost-covering *with* discrimination at least takes some consumers' surplus into account, it is for this reason a better guide to investment than, say, average cost pricing. For some of the consumers' surplus on the road will be brought into account by discriminatory charging, while all surpluses are ignored by average cost pricing. But it is impossible to say whether this is a good or a bad thing. On the one hand, since much investment is undertaken provided it will pay without the help of discrimination, we might think that discrimination would lead to over-investment in roads. On the other hand, since most investment ignores all losses and gains on related goods, there is no reason to think there is or is not anything optimum about the existing pattern of investment. It is then an open question whether discrimination or average cost pricing would give us road investment nearer to, or farther from, some imaginary optimum.

The rule that all costs must be covered by receipts not only guides investment, but also, within limits, determines output. Once a road has been built, users must be charged in such a way that total costs are covered. If a fine degree of discrimination is possible, there will be no need to restrict the use of the road below what it would be with marginal cost pricing. For those who are only willing to pay marginal cost can then be charged marginal cost, while those who will pay more, are charged more. But if, as is likely, not much discrimination is practicable, costs may only be recovered by charging everyone a price near to average cost, and thus cutting out the marginal travellers.

---

[1] J. R. Hicks, *Bulletin de L'Institut de Science Economique Appliquée*. N.2. December 1946. J. E. Meade, *Trade and Welfare*, Ch. VIII.

[2] This implies neither that how investment decisions are made is unimportant, nor that economic argument is irrelevant here.

Suppose we charge for the use of a particular road by toll. There are a limited number of ways in which in practice a toll could be varied between users. We might perhaps charge different amounts to different types of vehicles; more to large than small. We could charge by the number of people in each vehicle, or vary the charge according to how often the journey was made. But it is hard to see how a fine degree of discrimination could be achieved, and we might end up charging everyone much above marginal cost. In this case the use of the road might be restricted unduly. Many people prepared to pay the minute extra wear and tear which each individually would cause, might be prevented by the charge from travelling. Thus the rule that costs must be covered by receipts may result in people being needlessly kept off new roads by high charges.

A last objection to Mr. Munby's rule is that it may involve unfair charging. For it relies on discrimination wherever this is possible. Mr. Munby, following Mr. Lewis, writes as if there could be no objection to charging different people different amounts for identical physical services, more to those who need or want the service more, and less to those prepared to pay less. Thus those who can only get suitable work in London would pay more to travel on M.1 than those on the margin of doubt whether to stay in Birmingham. There *may* be cases in which one would not feel it unfair to charge more to those prepared to pay more, if, for example, they are richer, or they just have a fancy for the thing in question. It might possibly seem fair to charge someone a lot for driving on M.1 if he just had a fancy for bowling along in his car. But where being prepared to pay more arises from greater need for the service, it would seem to most people unfair to charge more. Why should those who have bought a house in the country and need to travel to work pay more for travelling at a given hour than someone who lives in the city and wants to drive out for a picnic? Should the deaf, if they can be made to, really pay the indivisible costs of television?

It is sometimes argued that charging is ' fair ' if users of a good between them pay its costs, but unfair if non-users have to subsidise them. This seems, however, to be supposing a community of interest which does not in fact exist, between different users of the same service. A man who is charged £10 for a journey while his neighbour travels for £1 has as good a right to complain as a vegetarian tax-payer who helps to subsidise beef. Thus while it may in some circumstances be no worse for the marginal road user to be subsidised by his fellow traveller than by the tax-payer, where willingness to pay more arises from greater need for the service it would seem unfair to put the burden on other road users.[1]

---

[1] Pigou held that it was in general more fair to subsidise than to make up for below

We will now turn to Mr. Lewis' arguments, the arguments on which Mr. Munby's rules rest. Mr. Lewis argues not that all costs must be covered by receipts, but that at least certain indivisible ones must be, those which he calls escapable.[1] His arguments are of two sorts, (though he himself does not recognise this). Some are to do with the allocation of resources; others are to do with fairness, with who ought to pay for what. We will begin with allocative considerations.

Here Mr. Lewis' main innovation is the concept of escapability. He tries to show that in order to secure optimum allocation, escapable costs must be covered by receipts. He divides costs as follows into escapable and inescapable. All costs are escapable except two classes. Firstly, the cost of renewable assets no longer wanted because demand has fallen, and secondly, such part of the costs of permanent, non-renewable assets not being used to capacity as is in excess of their value in some other use. We shall hardly be concerned with assets not wanted because demand has fallen. We can therefore think of escapable costs as including the costs of all assets which might need renewing, while the costs of (some) permanent assets are inescapable.

Lewis seeks to show that what are ordinarily considered fixed costs are more like variable ones than is usually thought. ' Hiring contracts and assets ' he says, ' are in the same category '.[2] Of both it is true that part of the cost can be immediately escaped, while if more time is allowed to elapse, what can ultimately be escaped usually exceeds what can be immediately escaped. Thus if a building is hired it may be possible immediately to transfer some of the cost (under a lease) but when the contract expires ' with that comes complete escape '. Similarly with assets purchased. User cost is immediately escapable, but when the asset wears out all its costs are escaped. Only the cost of certain assets which are never going to need replacing are inescapable. Buildings and weekly hired labour are thus alike; the costs of both are at least ultimately escapable, and in this respect they differ from the cost of acquiring land on which to build a road. Since this cost never has to be repeated, it can never be escaped.

By calling most of what are ordinarily known as fixed costs ' ultimately escapable ', and thus classing them along with variable costs as escapable, Mr. Lewis brings out a resemblance, and conceals a difference. Both station platforms and weekly hired labour have to be rebuilt or rehired sometime. This is an important resemblance between them. It is plainly necessary to have a way of deciding whether to renew platforms, just as it is necessary to have a way of deciding how much labour to use on the platform when it has been built. If this were all that were meant by calling both escapable, no one could object. But this terminology conceals a difference. Hired labour can

---

cost charging by discrimination. Where, on the railways, there is a case for charging some people below cost he held that ' There seems to be no good reason for throwing this burden upon persons who make use of the service of railways rather than upon the general body of tax-payers.' *Economics of Welfare*, p. 314.

[1] Lewis' rules resemble most closely Mr. Munby's rule for under-utilised roads.

[2] Op. cit., p. 11.

for the most part be varied finely with output, while station platforms cannot. Once the platform is built, innumerable decisions are still possible about what labour and other more variable factors to use with it. It may be, as we have seen, that a use of the price mechanism quite well suited to decisions about building platforms is sometimes less well suited to decisions about what use to make of them once they are built. Profitability may be a reasonable guide to building them, but sometimes unduly restrict their use after they have been built. We can agree with Mr. Lewis that some way of deciding when to renew platforms is needed, without agreeing that profitability is always the best method to adopt.

Mr. Lewis argues that since it is obvious that receipts should cover immediately escapable indivisibles it therefore follows that they should also cover ultimately escapable ones.

'Indivisible costs,' he says, 'may be immediately escapable. A railway can decide to run more or fewer carriages, or trains, or to keep open more or fewer stations, or, taking the matter to its logical conclusion, to close down altogether. Whatever the indivisible unit, if the expenses attributable to it are escapable, they must be covered if the unit is to be maintained. Confusion on this point has been due to a tacit acceptance of the unit of product as being the unit on which the calculation of economic cost must be based. But there is no greater significance attaching to the passenger in transport than there is to one or other of the indivisible units required for his conveyance . . .[1]

The first cost mentioned is that of running more or fewer carriages. Mr. Lewis classes this as an immediately escapable *indivisible* cost presumably because a carriage can hold more than one passenger. It is arguable, however, that in any practicable system of marginal cost pricing, the costs of the last carriage would have to be averaged over its passengers. One could therefore consider the running costs of the last carriage as part of marginal cost, and argue that unless those who use it are prepared between them to pay the running costs, the carriage should not be attached to the train. Our readiness to agree that passengers should at least pay running costs of the last carriage is probably due to the fact that we do think of such costs in this way. When we pass on to the next case, the running of more or fewer trains, it becomes clear that immediate escapability is not a conclusive ground for requiring that an indivisible cost be covered by receipts.

The cost of running a train from X to Y is a technical indivisibility. A train holds so many people that it is not reasonable to assume that on the whole everyone travelling would be prepared to pay the same price for the journey. A large part of the cost is engine cost, and this must be incurred for one carriage, and rises comparatively little as more carriages are added. Thus a price which would just cover the extra costs of attaching another carriage, would not cover engine costs. There is then a possible choice between charging a price which will cover engine costs and so help decide whether the engine should run, or charging a price which will cover only what for purposes of illustration we may call marginal carriage costs, which will help decide how many carriages to attach. If the price covers engine costs, we may attach too few carriages, if it covers only carriage cost we may have difficulty deciding whether the engine should run.

---

[1] Op. cit., p. 13.

In practice there are so many joint costs in railways which present difficulties for allocation and pricing that we might well feel at least engine costs, which are clearly allocable, had better be covered by receipts. But this is a judgment not forced on us by economic principles. That an indivisibility is immediately escapable does not mean that we get optimum allocation by using profitability as a way of deciding on its use.

The last examples in the passage from Mr. Lewis refer to costs which may not be escapable for a long time, such as station buildings and platforms. Here the choice between using price for variable factors only, and using it via profitability for variable and fixed, could become a serious one.

Suppose that station buildings and platforms have several decades of life yet to run, and are not being used to capacity. We could follow Mr. Lewis, and charge prices that will cover the use of these facilities in order, when the time comes for renewal, to see whether they have been profitable in the past. But we could equally well let their costs go uncovered, in order perhaps that they shall be more fully used in their lifetime, judging that notional calculations will enable us to decide whether to renew them when the time comes.

In the passage we have quoted Mr. Lewis passes from cases of immediately escapable cost to others where costs are ultimately escapable. He invites us to agree with him that just as passengers ought to pay at least running costs of the marginal carriage, so, for they are escapable too, station buildings and platforms should never go uncovered. However, the real reason why the running costs of the marginal carriage should be covered is not that they are immediately escapable, but that for all practical purposes they are a part of marginal cost. When we come to indivisibles like the running of a train, although they are immediately escapable it cannot be argued that covering these costs by receipts is a necessary condition for optimum allocation. A fortiori, it cannot be so argued for ultimate escapables like buildings.

At the end of the passage quoted above, it appears that Mr. Lewis sometimes feels profitability is the only possible way of allocating indivisibles. If we do not agree with him that station platform costs must always be covered by receipts, this just shows, he says, that we fail to understand that ‘ there is no greater significance attaching to the passenger in transport than there is to one or other of the indivisible units used for his conveyance ’. It is not clear how one compares the significance of a man and a station; what seems to be meant, however, is that it is as important to decide wisely how many stations to build as how many people shall use them. This is indisputable. But it is strange that one as happy among surplus calculations as Mr. Lewis should feel that profitability is the only possible way of deciding when to build, or rebuild, a station.

Let us contrast Mr. Lewis' division between escapable and inescapable costs with the conventional distinction between fixed and variable. It is usual to put costs which can be varied quickly and continuously with output into one box—variable; while fixed include both costs of assets which last a long time but require eventual renewal and those which are permanent. Mr. Lewis seeks to show that the usual classification is arbitrary, and draws the line between those which are permanent and those which require eventual renewal.[1] But this classification can seem arbitrary too. Suppose an electricity main already laid would last 50 years. Consumers must not be attached to it, says Lewis, unless between them they pay its costs. If they would pay for all the costs of electricity except that of the main, they must nevertheless not have the electricity. On no account must there be 'uncovered escapable indivisible expenses which ought to be covered if the retention of these resources by the consumers of this commodity is to be justified'. If, however, the main happens to be made of some permanent substance, and its costs cannot be covered by receipts, it is quite all right for consumers to be attached to it and have their electricity consumption subsidized to the extent of its costs. In the phrase 'retention of these resources (the resources used to build the main) by the consumers' we have a picture of the price mechanism guiding not only divisible but indivisible resources. But if the question of 'retaining' the resources is not going to come up till the main wears out 50 years hence, we may feel it better to get the best use of the existing main meanwhile, even at the expense of increased difficulty in deciding whether to renew it when the time comes.

Mr. Lewis is concerned not only with allocation of resources but also with fairness between groups of consumers and producers. He holds that it is right and proper that certain costs should be paid by certain groups of people. He is more concerned about who pays for what goods, than about consumers' relative real incomes.[2]

The rule that receipts must always cover escapable indivisible costs, if possible by discrimination, is supported both by references to allocation, and by moral approval of discrimination. 'Where there are escapable indivisible expenses to be covered, the case for discrimination is clear. It secures an output nearer the optimum, and levies the indivisible cost on those who get the greatest benefit (measured by their consumers' surplus) from retaining the indivisible resource in this line of production.'[3] Let us examine the moral principle that those who get the greatest benefit should pay the most. There is a sense of 'getting the greatest benefit' in which this principle is plausible. If two people use a public facility, but one uses it more than the other, this is *a* reason (though not necessarily a conclusive one) for charging more to the one who uses it most. Other things being equal, it would seem fair to charge those who use public parks a lot, more than those who use them a little;

---

[1] This ignores his treatment of assets no longer wanted because of a fall in demand.

[2] I. M. D. Little, op. cit., Ch. XI.

[3] Lewis, op. cit., p. 21.

those who travel a long way on public transport should in fairness pay more than those whose journeys are short, and so on. ‘ He who benefits most ’ in this sense simply consumes more service than he who benefits least.

We saw earlier, however, that readiness to pay more for the *same* service may be due to greater need of it. A man with a large family may be prepared to pay more for food or fuel than one with a small family. In such cases ‘ the case for discrimination ’ is far from clear.

These arguments about who ought in fairness to pay what, are importantly different from arguments about what price ought to be charged in order to secure optimum allocation of resources. In argument about optimum allocation price is merely the means of deciding what output of each good is to be produced. In these ethical arguments price is not being considered as a tool for influencing output, but as a means of dividing costs between different groups of people. Failure to distinguish these different types of argument sometimes makes Mr. Lewis' reasoning hard to follow. Let us take as an instance one paragraph designed to show that escapable indivisible expenses ought to be covered by receipts. (Ought in fairness? Or ought, as a condition for optimum allocation?)

‘ . . . The rule that price must equal marginal cost is inadequate; the rule should rather be that price must be not less than marginal cost, and there must be enough surplus to cover escapable indivisible expenses. This is obvious enough if the indivisible cost is a ‘ customer cost ’ incurred for one consumer only, e.g. the cost of connecting his house for electricity. . . . But it applies equally if the indivisible cost benefits a number of consumers, e.g. the laying of a main for the whole street, or all the consumers ’.[1] Here we cannot quite tell whether an allocative principle or an ethical one supports his statements. When he says it is obvious that a customer ought to pay his own connection costs, is it obvious because it is fair for each to pay the cost of what he consumes, or because for him not to pay these costs will lead to misallocation? Neither of these grounds, when separated out, can be extended to show that consumers' between them ought to pay somehow for a joint indivisibility.

Suppose that the reason why X ought to pay his own connection cost is that it is fair for every consumer to pay the cost of his own consumption. If then X and Y share the same main, we cannot infer that X merely because he is willing, should pay more of the cost of it than Y. Suppose on the other hand Mr. Lewis is really arguing that X ought to pay his own connection costs because if he is *unwilling* to do so it is better for the resources involved to be used in some other way. It does not follow that the best use of resources is made by consumers between them paying the costs of the main. They might be thought *willing* between them to pay for it by discrimination, but the charging system required be impossible to administer. Thus though a strong case for a customer paying his own connection costs can be made out both on grounds of fairness, and in the interests of good allocation, neither of these grounds can be extended to support the claim that a main must always be paid for by just those persons attached to it. It is only when the ethical and allocative arguments are not distinguished that the extension seems justified.

---

[1] Lewis, op. cit., p. 20.

Mr. Lewis' arguments to show that certain indivisibles ought always to be covered by receipts can be justified neither by the theory of optimum allocation nor by acceptable moral principles. It follows that Mr. Munby's rule for new roads also lacks justification.

St. Hilda's College,
Oxford.

SOME PROBLEMS OF PLANNING THE IMPROVEMENT OF THE ROAD SYSTEM

By D. J. Reynolds

The aim in planning the improvement of the road system and of assessing priorities is to obtain the road system which, bearing in mind the costs of achieving it, best reflects the wants and needs of the community for road space.

The planning of the improvement of the road system involves a survey of the existing road system on a uniform basis against realistic standards and criteria to discover the stretches of road where improvement is most needed and the type and scale of improvement necessary at each point. Such a survey would also indicate the expenditure required to bring the road system up to a desirable standard. More detailed analysis of the improvements suggested by the survey and the selection of those improvements giving the maximum benefits for the resources currently devoted to road improvement may then enable priorities to be correctly assessed.

Eventually it is hoped to develop techniques and methods whereby the assessment of road needs, of priorities and of the funds to be devoted to road improvement may all be carried out as an integrated process since they are all merely different aspects of the same thing i.e. the allocation and direction of resources to satisfy the community's wants and needs for road transport.

This paper discusses some of the general problems of road planning and puts forward purely for consideration some general principles and techniques which might help to deal with them. Obviously such a study can be extremely wide-ranging and complex and initially many simplifying assumptions and omissions must be made. Thus because of its complexity and because of the many possible alternative approaches and solutions to urban traffic problems attention will be confined to the planning of rural roads, where it may be assumed that the solution is fairly unambiguously one of road improvement and where existing traffic flows may be taken as data, and the diversion and generation of traffic (or the expansion in demand) ignored, with much less risk of error.

Again it will be assumed that the effects of road improvement in the speeding up of traffic, the shortening of distances and the reduction in accidents are objectively measurable in money terms and reasonably comparable with the cost of road improvement schemes.

As indicated above this analysis will be largely confined to the general principles which may be of help in this process. The detailed standards and criteria of width, alignment, etc., referred to above, will not be considered here therefore. The obvious starting point for the analysis is the existing rural road system and the existing traffic flows on it.

(i) *Investment Problems.*

Most road improvement schemes involve investment and the sacrifice of current resources in order to obtain more efficient road facilities at a later date, the investment yielding up its savings over the subsequent period. Assuming that there is a rate of interest (I) which genuinely expresses persons' preferences for present rather than future resources such that the sum of £(1 + I) in one year's time is only worth £1 now, we have to compare the present (capital) value (V) of a stream of net benefits of B per annum with the capital cost of improvement C.

Now the capital value of a stream of benefits extending indefinitely into the future is given by:

$$V = \frac{B}{1+I} + \frac{B}{(1+I)^2} + \frac{B}{(1+I)^3} + \ldots \frac{B}{(1+i)^n} \qquad (1)$$

where n is very large.

Multiplying both sides by $1 + I$ it can be shown that:

$$V = \frac{B}{I} \qquad (2)$$

For the improvement to be worth while (prima facie) $\frac{B}{I} > C$ or $\frac{B}{CI} > 1$ (the basis for the benefit/cost ratio method of comparison), or alternatively, $\frac{B}{C} > I$ (the basis of the rate of return method of assessment).

In reality however the situation is not as simple as this. In the first place it can be no means certain that the stream of net benefits will continue indefinitely into the future for the road may be superseded eventually by an alternative route, or road transport may possibly be superseded or at least diminished by alternative means of transport in the future. To deal with this possibility therefore it is suggested that benefits should be discounted at a somewhat higher rate than the market rate of interest.[1]

Thus if for example the road were to be superseded in 20 years, the rate of interest would need to be 8 per cent rather than an assumed market rate of 5 per cent.

In the second place the above example implicitly assumes that sufficient investment funds will be forthcoming to carry out all road investments that are prima facie worthwhile i.e. would yield a rate of return at least equal to the rate of interest. In practice however there is probably a back-log of worthwhile road improvements which investments funds will not be sufficient to carry out for some time, even though these funds may be eventually increased to the (correct) annual level indicated by a road 'needs' study. In this situation the question tends to become one of obtaining the maximum benefit from the resources currently devoted to road investment, and selecting the most worthwhile improvements which will just exhaust the available resources, with the rate of return on the least worthwhile investment which can be carried out indicating whether or not a particular improvement is worthwhile. This again suggests the use of a higher rate of return or a higher rate of interest than the market rates for planning purposes, although some consideration must be given to the possibility that investment funds may increase in the future sufficiently to plan road improvements on a more generous scale, and in these circumstances it would be a mistake to plan solely on the basis of the returns available on the most profitable present schemes.

---

[1] This seems preferable to depreciating the improvement over some conventional or arbitrary 'life', because provided that the improvement is maintained intact, and net benefits adjusted for these maintenance costs accordingly, the improved road may have an almost infinitely long *service* life.

In the third place most road investments are undertaken against a background of rapid increase in traffic which may be expected to continue into the future. Since gains from improvements tend to increase at least as rapidly as traffic flow, the discounted annual benefits are quite likely to increase in value with time for a number of years. For example, if the simple assumption is made that the costs per vehicle mile borne by road users are a constant plus an additional cost that increases linearly with traffic flow and in inverse proportion to road width, then costs per vehicle mile will be given as $a + \frac{bq}{W}$ where a and b are constants, q=traffic flow in vehicles per period and W= road width. Then the benefits per period from an increment in road width w will be $\frac{wbq^2}{W(W+w)}$ i.e. benefits will increase with the square of traffic flow and discounted benefits would become very large if traffic continued to increase, to which of course there would be a limit in practice. This problem is sometimes dealt with by assuming diminishing rates of increase or some ceiling to traffic at some future date but such methods may be complicated and laborious and may be arbitrary. To a considerable extent however the difficulty of dealing with future increase in traffic is more apparent than real, for in fact the problem is merely one of keeping road widths etc. in step with an increasing traffic flow, rather than the virtually impossible task of choosing a single width or road facility which will give minimum (present) costs over all future time and traffic. Thus if traffic is increasing and a just worth-while annual increment in width of w (yielding benefits of B) is made to keep pace with traffic and maintain optimum road width, then the annual benefits from any past increment in width w will remain at B because in spite of rising traffic, successive increments will keep down the net benefits from any previous increment to B. Provided therefore that a policy of continuous improvement when required is to be carried out investments can be assessed and ranked in terms of immediate returns only and most of the problems connected with investment policy to cope with increasing traffic volumes disappear.

They do not disappear completely however because of the question of indivisibility and the optimum size of the increments to be added to roads and the road system. This question has two aspects. In general fixed costs irrespective of the extent of improvement are incurred whenever a road is improved. These may be planning or design costs or they may be fixed costs of construction such as the utilisation of plant or costly alterations in buildings and structures in order to change road width. The very frequent improvement of roads in small increments to keep closely in step with traffic flow is therefore likely to be too costly a policy and the most economical policy will be one of larger less frequent improvements. The second aspect is that for rural roads at least it is conventional, and presumably more efficient for traffic, to build and expand roads in units of one lane and this again indicates some indivisibility in road investment. This indivisibility, besides raising problems of its own, weakens the validity of the assumption of continuous improvement to keep closely in step with traffic, but does not destroy it; it merely means that road facilities will not keep precisely in step but will first lag behind and then jump ahead of traffic increase (see below).

(ii) *Indivisibility and Looking Ahead*

As indicated above this arises from fixed costs of road improvement and from the greater efficiency, for rural roads at least, of having roads with definite lanes. The former aspect is considered on a highly simplified basis in the Appendix from which it may be seen that if it costs the constant sum of F per mile in fixed costs in order to change road width and R per mile in variable construction costs in order to add one foot to road width, then the road should be widened to accommodate an increasing traffic flow when:

$$Wo = W_1\left(1 + \sqrt{\frac{F}{RW_1}}\right) \quad \text{or} \quad Wo - W_1 = \sqrt{\frac{FW_1}{R}} \qquad (1)$$

where Wo = width appropriate to present traffic flow on basis of variable costs of widening only and W_1 = original road width.

If traffic flow is expected to continue to increase and the fixed costs of changing road width are constant at F for future changes in road width then the most economical increment in road width will be given approximately by:

$$W_2 - W_1 = 2\sqrt{\frac{FWo}{R}} \qquad (2)$$

where W_2 = new road width, W_1 = original road width, Wo = road width, appropriate to present traffic flow on basis of variable costs only.

This indicates that with constant fixed costs of changing road width width should first lag behind traffic and then jump ahead of it by about the same amount when it becomes worthwhile improving the road. If however the next widening is likely to cost more or less in fixed costs then the present widening then the most economical increment in road width will be given (again approximately) by:

$$W_2 - W_1 = \sqrt{\frac{F_1Wo}{R}} + \sqrt{\frac{F_2Wo}{R}} \qquad (3)$$

where F_1=fixed costs incurred in present change in road width and F_2=fixed costs likely to be incurred in next change in road width.

The significance of these relationships lies in the fact that the question of planning road improvement ahead to cater for future traffic increase seems to be almost entirely determined by the fixed and variable costs of changing road width. Future traffic levels and rates of increase in traffic seem to play little direct part, a high rate of increase for example being mainly accommodated by more frequent inprovements. Nor does there seem to be any specified period to look ahead when increasing the capacity of roads, although allowance must of course be made in the design for traffic increase during the period required for the planning and carrying out of improvements.

This aspect of indivisibility is probably most important in urban areas where the other aspect of indivisibility (lanes) is less important and where fixed costs in demolishing buildings, etc., may be so high in relation to variable costs that widenings may be only worthwhile when road width has lagged considerably behind traffic with a corresponding large leap ahead when widening is undertaken. In rural areas fixed costs of improvement in relation to variable are probably much lower and the lane aspect of indivisibility tends to take over and become the more important. Lanes are then usually added when they become worthwhile i.e. when the total (current) costs of the road plus those of the given volume of traffic are the same on a 3 lane road for example as on a 2 lane road.

Even in rural areas the fixed cost aspect of indivisibility can be important however because this aspect applies not only to the whole road but to its various component parts. Thus if the next widening is likely to be very costly because of likely building developments approaching close to the carriageway, it will be worth while acquiring land in considerable excess of immediate requirements in order to cover this eventuality, or if structures such as bridges are built which would be relatively expensive to alter and widen it will be worthwhile building them in considerable excess of immediate requirements. An interesting special case of fixed costs in rural improvements occurs in the normal transition from a 2 to a 3 lane and then to a 2 x 2 lane dual carriageway as traffic expands [1]. Traffic increase on a 2 lane road would normally suggest conversion to 3 lane and to dual carriageway later if required. The transition from 3 lane to dual carriageway will normally be costly however for part of the 3 lane carriageway will need to be converted to a central reservation and the second carriageway added, or if space is available, the central reservation and the second carriageway may be added to the 3 lane road giving an asymmetrical road with one lane not fully required. In either case the cost of ~~the extr~~a lane added will in effect be the cost of construction of two lanes, with the cost of one lane as the fixed cost of the change. This fixed cost of change is so high that the most economical increment in width given in (3) above may be a direct transition from a single 2 lane to a 2 x 2 lane dual carriageway particularly as the latter is likely to have the advantage in safety. Thus it may be noted from (3) above that using foot-widths of carriageway as units of cost and putting $R=1$ with $Wo=33$ ft. (a 3 lane carriageway) $F_1=1$ and $F_2=11$ (the width of a lane) the most economical increment in width is in fact just under 25 feet (2 lanes).

(iii) *Diminishing Returns*

Diminishing returns to successive increments of improvement at any one point and time and traffic flow are likely to arise in many cases simply because on normal traffic flow/cost relationships the gains from the addition of the last foot (or lane) of road width will be less than that from the penultimate foot (or lane) of width. Thus it is shown by the simplified traffic flow/cost relationship given above that returns tend to diminish in inverse proportion to the square of road width. On a more realistic basis it may be assumed that costs of traffic per vehicle mile will vary linearly with journey time per mile so that average cost per vehicle mile will be given by the form $a+\frac{b}{V}$ where a and b are constants and V=average speed in m.p.h. A speed/flow relationship established by Wardrop [2] is in the form of a constant speed which diminishes linearly with traffic flow in inverse proportion to road width. Thus:

$$V=c-\frac{d(q+x)}{W-y}$$

where V=average speed in m.p.h., c, d, x and y are constants and q=traffic in vehicles per period and W=road width in feet.

Then the total savings to traffic per period from a unit increase in road width will be given by:

$$\frac{q\,b\,d\,(q+x)}{(c(W\text{-}y)-d(q+x))^2}.$$

That is returns still tend to diminish (but more sharply) in inverse proportion to the square of road width.

Diminishing returns to the (lateral) scale of improvement have important implications for road planning because they suggest that depending on the profitability of the various investment schemes the greatest return from road investment will be obtained from the spreading of investment widely and thinly over the stretches of road where improvement is worthwhile particularly as this may appear to be an equitable kind of distribution for centrally allocated funds. This simple view ignores however the fixed costs of improvement which must be incurred irrespective of the extent of improvement and these fixed costs alone would suggest some greater concentration of road investment. It can be shown in fact (see Appendix) that with a fixed cost of (present) widening of F_1 per mile and variable costs of R per mile that the overall rate of return on a widening will tend to increase up to the point that the width added $=\sqrt{\frac{F_1W_1}{R}}$ where W_1 original road width.

However fixed costs of improvement are probably low in rural areas and although the lane aspect of indivisibility again tends to diminish the significance of diminishing returns, it is still likely to be of some importance in rural road planning.

Diminishing returns to the scale of improvement mean that returns from a given road investment are determined, not only by the original road problem and the type of improvement selected to deal with it, but also by the extent to which the improvement is carried. Thus for a particular road problem an improvement A may show an overall rate of return of 6 per cent made up of equally costly increments giving returns of 7 per cent, 6 per cent and 5 per cent, whereas an improvement B may show an overall rate of return of 5 per cent made up of equally costly increments giving 8 per cent, 6 per cent, 4 per cent and 2 per cent. It might be concluded therefore that A is the better solution to the problem, whereas if each improvement had been carried to a common and appropriate margin, say 6 per cent, it would be seen that solution B with an overall rate of return of 7 per cent compared with 6½ per cent for A is the more worthwhile. The original rate of return for B is the lower simply because it had been carried the further. Thus the overall rates of return or benefit/cost ratios on improvement schemes are not necessarily an accurate indicator of the relative worthwhileness or profitability of such improvement schemes; they are only an accurate indicator when each improvement is carried to a common or appropriate margin.

In particular when the rate of return on a particular scheme is high it must be asked whether it is high because of the inherent profitability of the situation and the scheme selected to deal with it or merely because the improvement has not been carried far enough.

Similarly if the rate of return on a scheme is low this may be because the situation and scheme is inherently unprofitable or it may be because it has been carried too far, much of the expenditure on the scheme having been incurred for little immediate benefit.

Although it seems clear that, with the limitations given above, diminishing returns apply to road widenings or the width to which new facilities are constructed, it is not clear how generally they apply to road improvement schemes. For example when replacing a curved section of road by a straighter section of road it is clear that increasing returns may well apply, for gains will tend to increase and the costs tend to diminish the more the road is straightened. It would be unwise therefore to assume that diminishing returns to road improvement schemes are universal.

(iv) *Discrimination and the Marginal Approach*

The term discrimination is used to indicate the splitting up and consideration of the road system in reasonably homogenous parts with appropriate remedies for each. This is preferable to an indiscriminate approach aimed at common overall consideration and treatment of whole stretches of road containing sections of widely differing characteristics and conditions.

This suggests the division of the existing road system into homogeneous or nearly homogeneous parts as far as possible with appropriate treatment and remedies for each part and this in turn suggests a ' black spot ' or ' bottleneck ' type of approach to road improvement, the first candidates for improvement naturally being those showing the greatest rates of return. In common with many general principles however discrmination cannot be carried too far because in the first place such an approach might well produce a road system of rapidly changing widths and other characteristics with obvious disadvantages; there are in fact advantages in uniformity of the road system for its own sake.

In the second place it is necessary to extend the principle of discrimination beyond short to longer stretches of road in order to ensure that the best improvement is carried out. For example one may have the situation where a badly aligned stretch of road AB is contained within a longer stretch of road CD. Narrow discrimination would suggest the straightening of the length AB but extension of the principle of discrimination might well indicate that the most worthwhile investment would be a new stretch of road running direct between C and D, which might make the straightening of section AB superfluous.

An essential part of the discrimination process therefore is the marginal approach and the comparison of benefits with costs for successive increments which has been outlined above in connection with road width and diminishing returns. Applied to the (longitudinal) extent of improvements it simply means starting with the section of road which it is most profitable to improve and then extending the improvement taking account of the marginal increase (or increment) in net benefits vis à vis the marginal increase (or increment) in costs for each step in the extension finally stopping when marginal benefits equal marginal costs and no further net savings can be obtained. Thus for the above example one would first consider the benefits and costs of straightening section AB of the road and then go on to consider the increment in benefits and the increment in costs to be obtained from installing a new road between C and D. If the increment in benefits exceeds the increment in costs then a new road between C and D would be the desirable extent of improvement and if the majority of traffic deserts the old road the re-alignment of section AB would not be worthwhile.

The use of discrimination and of the marginal approach indicates a piecemeal and gradual policy for road investment the scope of schemes being gradually extended with increments of costs and benefits being compared at each stage. Thus the size and scope of improvements are gradually built up stage by stage rather than choosing in the first instance a comprehensive scheme to deal with many stretches of road of widely differing characteristics and problems. When considering a whole stretch of trunk road with the common pattern of more or less adequate stretches of road through open country with short stretches through built-up areas, discrimination and a marginal approach would suggest consideration of first, by-passes for the built-up areas, second, improvement to the existing sections of rural road and finally the construction of a motorway to by-pass the whole length of road. Increments of costs and of benefits should be calculated for each stage. Discrimination does not, of course, necessarily preclude comprehensive improvements of the motorway type but merely insists that their additional costs over smaller improvements must be justified by their additional benefits.

(v) *Economies of Scale.*

The term economies of scale is used to describe the fact that because of the nature of traffic and of flow/speed relationships, a greater traffic flow can be accommodated at a constant speed with a less than proportionate increase in road width e.g. an increase in traffic flow of 100 per cent can be accommodated by a less than 100 per cent increase in road width.

Thus with a rural flow/speed relationship of the form[1]

$$V=50.5-\frac{q+1380}{11(W-13)}$$

where V=average speed in m.p.h. q=traffic flow in vehicles per hour and W=road width in feet with a minimum value of 24 feet, it may be calculated that at an average speed of 35 m.p.h., about 500 vehicles per hour can be accommodated on a 2 lane (24 feet) road, about 2000 vehicles per hour could be accommodated on a 3 lane (33 feet) road and about 5,000 vehicles per hour could be accommodated on a 4 lane (2 x 22 feet) road.[2] Although an overall formula of this kind is liable to considerable error there is no doubt that as width is increased the capacity of a road, in terms of the traffic volumes that can be accommodated at a constant speed, increases in much greater proportion than width.

This means that (other things being equal) a given traffic flow is much more economically accommodated on a single road of appropriate width than on say two separate roads of appropriate width or alternatively that there are considerable net advantages to be obtained by attracting and concentrating traffic on relatively few highways of high capacity even though some additional vehicle mileage may be necessary for traffic to carry out its journeys. This is especially true because the concentration of traffic onto a few highways will encourage the installation of facilities which are only worthwhile when large traffic volumes are catered for e.g. dual carriageways, fly-over junctions lighting, lay-bys, general safety measures, etc.

Economies of scale therefore point to the concentration of investment on major highways and on the existing trunk road system and the attraction of traffic to these routes, rather than the wider dispersal of investment over major and minor routes alike.

---

[1] See [3] in the list of references at the end of this article.

[2] It may be noted that the limiting capacities of rural roads recommended by the Ministry of Transport at present are as follows:

No. of lanes	Passenger-car units per 16-hour day
2	up to 4,500
3	4,500 to 9,000
2×2	9,000 to 25,000
2×3	over 25,000

It should perhaps be noted that economies of scale arise from the scale effect of greater traffic volumes and the (more economical) provision of road facilities to accommodate them. They do not conflict therefore with diminishing returns to road width, which arise from an increase in road width with traffic remaining constant e.g. from a ratio effect. Diminishing returns to road width and economies of scale are quite compatible with each other therefore on one and the same stretch of road.

(vi) *Summing Up and General Methodology in Planning the Improvement of the Road System*

It may be seen from the above that road improvement policy seems likely to be the outcome of several different principles and relationships, some of which are likely to conflict, rather than of a single principle or a consistent set of principles. In general these relationships and principles may be grouped into those favouring concentration of investment i.e. indivisibility and economies of scale, and those favouring widespread dispersal of investment i.e. diminishing returns and discrimination, but because of these conflicts the planning of the road system is likely to be a somewhat complex process and it is important that the general methods used should take account of this.

The planning of the road system would be concerned with recording the width alignment (horizontal and vertical) and general traffic conditions of the existing road system and the existing traffic flows on it, dividing it into homogeneous or near homogeneous sections as far as possible and then comparing each section with various standards of width, alignment, etc., or by using various techniques to decide in the light of costs of the various schemes whether improvement is required and whether widening, re-alignment, bypasses, etc., or a new road or motorway is the most worthwhile remedy.

It is obvious that this process of discriminate step by step planning of improvements and the possible number of alternative types and scale of improvement to solve a particular problem to be considered, can be a very laborious process and a heavy burden on scarce planning and design resources. Fortunately however the appropriate type or scale of improvement to deal with a particular problem may be quite unambiguous in many cases or will fall within a narrow range of alternatives most of which may be considered and rejected with little work. Again the fact that traffic is heavily concentrated on a small part of road mileage will tend to lessen the work involved and probably a survey of the rural trunk road system (5 per cent of total rural road mileage but carrying about 40 per cent of total rural vehicle mileage), [1] and of Class I roads (10 per cent of rural road mileage but carrying about 30 per cent of total rural vehicle mileage) will uncover the vast majority of worthwhile improvement schemes.

Nevertheless the survey and planning work is likely to be heavy and a considerable strain on planning and design resources and it is important therefore that these should be deployed in the most efficient way. Broadly

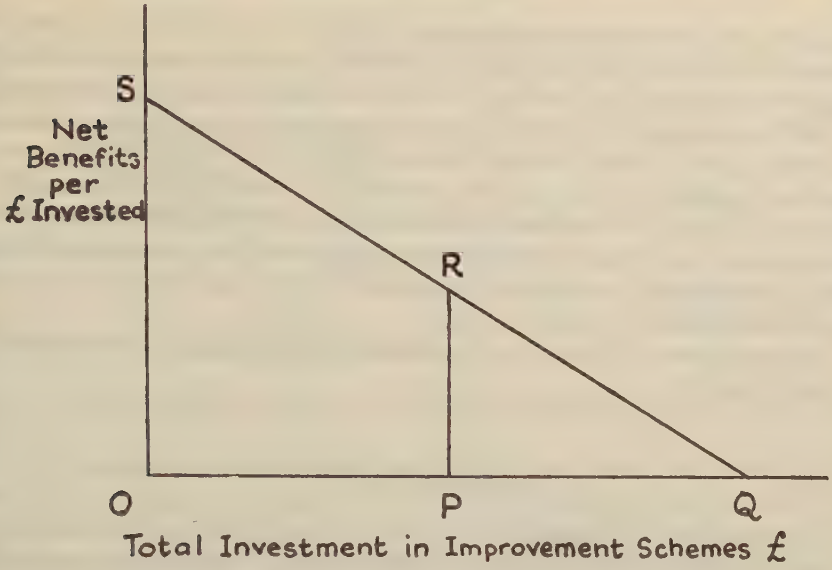


FIG. 1. Gains from a hypothetical road investment programme.

speaking it seems preferable to roughly survey the whole main road system using fairly simple methods and approximations and then gradually to increase the detail and accuracy of the study on the stretches of road where improvement seems to be worthwhile on the preliminary survey. This seems to be a more useful disposition of planning and design resources than a more narrow and arbitrary choice of problem sections of road but a more detailed initial study of these sections. This is especially true because in the nature of the road improvement problem, some schemes are likely to give much higher returns than others and it is much more important to survey the road system comprehensively to ensure that the high return schemes are included than to consider the details of design in particular cases. Thus the immediate returns from possible road improvement schemes in any one period may be arrayed in diminishing order as in Fig. 1 in which they are assumed to diminish linearly when so arrayed, due weight being given to the cost of individual schemes i.e. a scheme costing £2 occupies twice the length on the horizontal scale of a scheme costing £1.

Then if the sum of OP is available for investment in the period the most worthwhile schemes for that sum will give benefits of area OPRS, but if twice the sum of OP (=OQ) is available for investment benefits will only increase to $OQS = \frac{4}{3}OPRS$. In other words the particular schemes chosen to be carried out are much more important than the total sum of money allocated to road investment and the marginal details of design which are only likely to give marginal gains.

Many more details of methods to be used in planning the road system remain to be worked out. In particular, although the central allocation of funds for road investment suggests something like the general and abstract approach given here, there is a need to translate this into reasonable working rules and to devolve the planning of the road system in order to make use of the local knowledge and expertise which is essential in the planning process.

Although the whole planning task is likely to be rather difficult and laborious therefore, there can be little doubt that it would be well worthwhile, for rates of return on investment schemes can vary widely and even the comprehensive and lengthy investigation of the traffic and economics of the London-Birmingham motorway is estimated to have cost less than 0.1 per cent of the capital cost of that road [4].

Acknowledgment. This article is published by permission of the Director of Road Research.

REFERENCES

1 Charlesworth, G., Reynolds, D. J. and J. G. Wardrop. *Road Improvements: Choosing Priorities by a New Formula*. Engineering 1959, 188 (4873), 185–8.
2 Wardrop, J. G. 'Some theoretical aspects of road traffic research.' Proc. Instn. Civ. Engings. Pt. II, Vol. 1. June 1952.
3 Reynolds, D. J. *The Assessment of Priority for Road Improvements*. Road Research Technical Paper No. 48. H.M.S.O. 1960.
4 Coburn, T. M., Beesley, M. E. and D. J. Reynolds. *The London-Birmingham Motorway: Traffic and Economics*. Road Research Technical Paper No. 46. H.M.S.O., 1960.

Road Research Laboratory

APPENDIX

Fixed costs of improvement and the optimum increment in road width

To simplify the problem so as to illustrate the main factors involved it is assumed that the cost per vehicle mile of a given traffic flow will be given by

$$C = a + \frac{bq}{W}$$

where a and b are constants q=traffic flow in vehicles per hour and W=road width in feet.

Without fixed costs of changing road width and ignoring lanes the aim is simply to minimise the total costs (road costs plus costs to traffic) of accommodating the given traffic flow q which may be assumed to flow for H hours in the year. Ignoring road maintenance costs the aim is to minimise the annual sum:

$$\mathrm{WRI}+\mathrm{Hq}\left(a+\frac{bq}{W}\right)$$

where W = road width in feet.

R = capital cost of widening road by 1 foot for 1 mile.

I = annual rate of interest.

Differentiating with respect to road width (W) this is minimised when

$$\mathrm{dWRI}=\frac{\mathrm{dWHbq}^2}{W^2}.$$

That is when $W_0=q\sqrt{\frac{bH}{RI}}$.

When this condition is satisfied the road can be said to be at optimum width (W_0).

In this situation without fixed costs of changing width, the road is simply widened continuously to keep pace with an increasing traffic flow and no problem of size of improvements or of planning ahead rises. With fixed costs of changing width of F per mile however and with a road of given width W_1 and an increasing traffic flow a more complicated situation arises.

The road is then only just worth while widening when:

$$W_1RI+Hq\left(a+\frac{bq}{W_1}\right)=W_0RI+FI+Hq\left(a+\frac{bq}{W_0}\right)$$

where W_0 = the optimum width for present traffic flow q.

That is when

$$\frac{Hbq^2}{W_1}-\frac{Hbq^2}{W_0}=(W_0-W_1)RI+FI.$$

$$\therefore\quad \frac{Hbq^2(W_0-W_1)}{W_0W_1}=(W_0-W_1)\,RI+FI.$$

Now from (1) above $Hbq^2=W_0^2RI$ so that

$$RI\,(W_0-W_1)\left(\frac{W_0}{W_1}-1\right)=FI.$$

$$\therefore\quad (W_0-W_1)^2=\frac{FW_1}{R}\text{ and }W_0-W_1=\sqrt{\frac{FW_1}{R}}. \qquad (2)$$

Just as it is worth while lagging behind optimum road width until widening is sufficiently worth while to overcome fixed costs, so it is worth-while investing in excess of the width immediately required in order to reduce the frequency of future changes in width to accommodate an increasing traffic flow.

That is:

$$W_2RI+Hq\left(a+\frac{bq}{W_2}\right)=W_0RI+FI+Hq\left(a+\frac{bq}{W_0}\right)$$

where W_2 = final road width after widening.

$$(W_2—W_0)\ RI = \frac{Hbq^2}{W_0} - \frac{Hbq^2}{W_2} + FI.$$

$$\therefore\quad (W_2—W_0)\ RI = W_0^2 RI\left(\frac{W_2—W_0}{W_2W_0}\right) + FI.$$

$$\therefore\quad (W_2—W_0)^2 = \frac{FW_2}{R} \text{ and } W_2—W_0 = \sqrt{\frac{FW_2}{R}} \tag{3}$$

Summing (3) and (2) therefore the most economical increment in width will be given approximately by:

$$W_2—W_1 = 2\sqrt{\frac{FW_0}{R}}$$

or if the present widening costs F_1 and the next widening F_2 in fixed costs the most economical increment will be given approximately by:

$$W_2—W_1 = \sqrt{\frac{F_1W_1}{R}} + \sqrt{\frac{F_2W_0}{R}}$$

It should be noted that investment in the widening from W_1 to W_2 will not earn the full rate of interest I immediately because some of the investment is undertaken to reduce the future number of widenings and the immediate rate of return on the investment will in fact be

$$\frac{IR(W_2—W_1)}{R(W_2—W_1)+F}.$$

The maximum overall rate of return on a widening with fixed costs will be obtained when the marginal rate of return i.e. the rate of return on the last increment of width, is equal to the overall rate of return. That is when

$$\frac{Hbq^2}{RW_x^2} = \frac{Hbq^2(Wx—W_1)}{W_xW_1[R(W_x—W_1)+F_1]}.$$

where W_1 = original width and W_x = width at which overall rate of return on widening is maximised.

$$\therefore\quad RW_x(W_x—W_1) = RW_1(W_x—W_1) + F_1W_1.$$

$$\therefore\quad (W_x—W_1)^2 = \frac{F_1W_1}{R}$$

$$\therefore\quad W_x—W_1 = \sqrt{\frac{F_1W_1}{R}}.$$

THE ROAD RESEARCH LABORATORY'S INVESTMENT CRITERION EXAMINED

By D. W. Glassborow

Detailed calculations of economic return have not in the past usually been made in the case of new road construction or major improvement schemes in this country, but this year the Road Research Laboratory published the details of a study,[1] made in co-operation with the University of Birmingham, in which the expected economic and social benefits of the London-Birmingham motorway were calculated and related to its estimated capital cost.

The decision of principle to construct this motorway was, of course, taken in advance of the calculation of social return and made use of estimates of the capital cost which have turned out to be substantially too low. This does not, however, imply that the decision was a bad one, and many parallels could probably be found in both the private and the public sectors of the economy.

In this paper I do not wish to pursue the question whether the decision was in fact justified, though I suspect that equal benefits could have been achieved with a smaller expenditure, and that a substantial proportion of the expenditure could probably have been deferred, as the volume of traffic forecast for the first ten years is below the capacity of a two-lane dual carriageway. Furthermore, the study was a pioneer one in this country and it is natural that many minor details of the method and the estimates could be argued about. But there is an air of general statistical competence and skill in the use of economic reasoning about the paper. What I want to discuss is whether this method of calculating benefits is potentially a good one and if so how it should be used in road investment decisions.

This paper is therefore not concerned with the details of the calculations nor with the accuracy of the traffic forecasts—it is to be hoped that the Road Research Laboratory will soon publish some comprehensive figures of actual traffic comparable with the estimates for June–July 1960 given in the paper—nor with the accuracy of the estimates of capital cost—though it would be interesting and useful to know what the cost turned out to be when the bills came in. In this paper I consider first the validity of calculations of social benefit from an additional road investment, when the present road system is not being used efficiently, second the assessment of the value to the community of time saved, and third the validity of comparing the social return on capital arrived at by the M.1 method with the commercial or bankers' assessment of return on capital which is normally used in connection with investment decisions.

---

[1] *The London-Birmingham Motorway, Traffic and Economics.* Road Research Technical Paper No. 46. H.M.S.O. 1960.

The Case for Additional Road Capacity

Statements that it is essential to invest very large sums in the improvement of the British road system, when supported by arguments, are usually arrived at by the following kind of reasoning.

There are many roads where, because of the present volume of traffic, speeds are lower and more variable than they would be if there were less traffic or more road capacity. In consequence, the costs now borne by individual road users, both in terms of time consumed and of vehicle operating expenses, are higher than they would be with better roads. Increases in road capacity so that speeds could be increased with the present volume of traffic would, therefore, benefit the individual to the extent that he saves time or incurs lower operating expenses and this benefit justifies new road investment. It is further argued that substantial investment, in addition to that justified by the elimination of present traffic congestion, is required to be undertaken now to cope with the traffic demands which will be generated by the greatly increased numbers of vehicles which will be in use in the near future.

The increase in road capacity required by a given date in the future, and in consequence the total expenditure required in the interval, has usually been estimated from a comparison of existing roads with official design standards for roads in the case of rural roads. It is not clear whether the standards adopted are technical ones, based on the assumption that certain average speeds are desirable, or economic ones. If the latter it would be expected that the standards would differ according to the broad composition of the traffic, which determines the economic costs of a slower flow, and according to the cost of road improvement in the area. In the case of urban improvements the approach has been even more arbitrary. On a number of occasions broad assessments of the benefit to users, by valuing both time saved and operating costs reduced as the combined result of all these improvements have been made, and the aggregate benefits have been related to the total investment expenditure to show an annual rate of return to the community.

The results of such calculations are said to show that amounts of the order of £3,500 million ought to be spent in the next ten years or so on new road construction and major improvements to existing roads, in order to relieve road congestion delays. It has been suggested that the community would obtain benefits from such expenditure large enough to recoup it in a few years. The total sums estimated to be required and, equally, the benefits claimed are so large that the methods of assessment and the assumptions implicit in them require careful examination.

It is obvious that they should not unquestioningly be accepted, for the formulation of the official standards of road capacity in relation to traffic cannot be a matter of simple calculation, while there are, of course, even more formidable statistical and conceptual difficulties in practice in assessing the benefits. There is also the fundamental question whether the rate of return as calculated should be used by those taking road investment decisions as if it were identical with the commercial return which guides decisions about investment in the other productive sectors of the economy.

The Present Use of the Road System

From a superficial approach expenditure on improvements to the road system is desirable because of the overcrowding which occurs sometimes regularly, sometimes intermittently, both on roads inside towns and on roads leading to towns. Less obviously there are the benefits which can be obtained from increasing average journey speeds at present kept down by the nature of the roads themselves even where other vehicles are not obstructing movement.

There is implicit in calculations of the road investment required to raise journey speeds the assumption that it is out of the question to bring about a better use of the existing road system by changing the structure and incidence of taxation or by other state intervention in the working of the economy. Apart from the influence of direct and indirect taxation on the pattern of expenditure and of subsidy policies on the location of industry, there are also the effects of such matters as school hours, working hours, conventional holiday periods, rent restriction and town and country planning controls.

Among these factors affecting the use of roads, the structure and level of road taxation is one of the most important. With the present system of road taxation (which must be regarded partly as the price for road use and partly as one item in the general revenue tax structure), on the one hand most roads are under-utilised for some or all of the time while some roads are at times inadequate to meet demands. The natural approach to such a situation is first to ask whether the price is right or whether the road tax structure in fact aggravates the problems.

Those who advocate as necessary the large expenditure on road improvements quoted above no doubt believe that road users are paying in special road taxes much more than the current economic cost of providing roads. If this were so—the question is debatable—there would be no good economic reason for increasing the general level of road taxation, nor would altering the structure of road taxation make unnecessary a substantial programme of road construction. However, even then an alteration in the incidence of road taxation might well indicate that road construction is desirable in different places from those for which plans are at present being made.

The structure of road taxation is therefore of great importance, whether or not road users as a whole are paying the full economic costs of the roads which are provided. The cost of providing road space varies widely from place to place and from one time of the day, week or year to another, while the importance of the various cost elements making up the total vary significantly, particularly as between town and country. For example, the cost of the land used in the construction of M.1 is a very small part of its total cost, of which the latest estimate is £395,000 per mile. In the case of an urban road of the same capacity the cost of the land by itself could account for some 80 to 90 per cent of a total cost of £3,000,000 per mile or more.

These variations in cost would create no pricing problem if all vehicles were identical and if their use inevitably required the supply of a constant ' mix ' of roads of varying values, i.e., if for all journeys the proportion of the mileage run over motorways, country lanes, suburban streets and main roads in city centres, etc., and the proportion of peak use was the same. So long as the full price for the whole road system was paid by road users on a uniform basis according to distance, there would then be no need, in practice or in theory, for charging separately at different rates for the use of town roads on the one hand and of country roads on the other. Since no road user would then require town roads without country roads or *vice versa*, there would be no sense in trying to divide up the total payment between the two parts of the journey.

In real life, of course, vehicles do differ widely both in their requirements of road space and road construction and in the times and places of their use. It is therefore important to have a system of road taxation which, in combination with such other road use charges as are necessary and practicable, differentiates between road users much more in accordance with the real resources they use than does the present system. Appropriate contributions to general taxation can be considered separately. Currently, even allowing for the need for general tax revenue, the present system of taxation substantially overcharges some road users and undercharges others. This surely encourages a use of the present road system which is far from the best in the interests of the community, and alterations to the structure of road taxation would promote better use of the roads.

In consequence, while the present tax system continues, the incidence of road congestion cannot indicate where investment in road improvements is economically desirable and the calculation of the social benefits accruing from the forecast reduction of present congestion by such investment cannot provide an economic justification for the decision to build such roads.

However, the use of social return calculations of the M.1 type involves the implicit assumption that at present the choice made by road users brings about the optimum use of the present road system and that new investment adds to their choice and a new pattern of use emerges, which produces the optimum use of the improved system. If the positions both before and after the investment are not optimum positions, then the net benefit from the investment should not be assessed from the difference between them. The net benefit to the community as a result of the planned investment might thus be significantly different from that calculated. With a better road tax structure new investment in transport, including roads, would still probably be required, but the actual physical investment schemes to be undertaken would probably be very different, in their scale at least, from those at present planned. Greater benefits to the community would then follow than are likely to be achieved by the implementation of a road investment programme designed to eliminate road congestion without the aid of changes in the present road tax structure.

A first objection to the social return calculation therefore is that its basis is the present pattern of road use, which is obviously distorted because of the inefficiency of the present road tax system as an economic price structure for the use of roads. (Though it seems impossible to devise a perfect system of road taxation for this purpose, and the control of road use cannot be left entirely to market forces, improvements on the present tax system would be a powerful aid to physical regulation, reducing the amount of effort which at present is required. Any rationing system is made easier to apply, if the price mechanism is used to assist it, rather than to oppose it.)

The Valuation of Time Savings

A major assumption in the calculation of benefits is that time savings are equivalent to money benefits. In fact in the case of the London-Birmingham motorway calculations, the major part of the benefit calculated arises from this source. Accordingly, much turns on the valuation of time saved. There are two aspects to be considered: savings in working time and savings in leisure time.

So far as savings in working time are concerned, it has been assumed in the London-Birmingham motorway assessment that these should be valued at the hourly earnings before income tax of those concerned. It is difficult to see the justification for this.

Savings in working time are of two kinds. On the one hand there are savings in driving time by those whose work is driving and on the other there are savings in journey time by those whose work is done before and after the journey.

In both cases, of course, the effect of reducing journey time is to increase the man-hours available for other work or more leisure. The time saved has a value to the community equal to the marginal value of time to each individual, though the benefit of the time-saving may be shared between the employee, the employer and the employer's customer, in the first instance by higher hourly wages and shorter hours or in the second by higher profits and in the third by lower prices. Their shares will be determined by the bargaining situation, that is the interplay of supply and demand, which already has determined the value placed on time before the road improvement was made. How the time saved is employed, and how the benefits are shared does not, however, affect its valuation.

In the case of those whose work is driving, if the time saved is used for leisure it should be valued as additional leisure time. Because wages after all are payment for both time and effort leisure time should be valued at less than hourly wages after tax, the reward for time and effort combined. In practice, both the eagerness shown by workers to obtain their share of overtime, and the insistence in agreements to reduce working hours that the weekly wage shall be unchanged suggest that little value is placed on leisure and a high price on effort. Similarly, if the time saved is used in additional work, this will require additional effort and the time should again, I believe, be valued at the rate for leisure time, though it is possible to make a case that the community values additional working time more highly than the individual values additional leisure.

To the assessment of the value of time saved by those whose work is driving, there should be added an assessment of the value of effort saved in performing the journeys themselves. This might justify, in practice, the valuation of time saved at hourly wage earnings before tax to arrive at the benefit from saving in time and effort, but there is no *a priori* reason to assume that it is right to do so.

The latter argument, however, can hardly be accepted in the case of those travelling on business. It seems that in their case the time saved should be valued, at the most, at the value of additional leisure time, which will be less than their hourly income after tax. How much less is difficult in practice to assess. I should emphasise that it is not because it is marginal time that it should be valued at less than hourly wages or salaries, but because hourly wages and salaries represent the reward not only for giving up leisure but also for doing work. Additional leisure time can be valued at infinitesimally less than current leisure time.

The value of leisure time saved was assumed in the study to be between 4 and 8 shillings an hour to the individual. There are, of course, good reasons to doubt whether such high figures are justified. If they are to be judged by reference to earnings in working time, it is once more relevant to remember that these are taxed and are payment for time and effort combined.

There are thus reasons for suggesting that, as a result of an overvaluation of time savings, the London-Birmingham motorway assessment considerably overstates the value to the community of this investment.

The Use of the Social Return as a Guide to Investment

There is a third consideration which is of a quite different nature and which in practice is of even greater importance. This derives from the presentation of the results in the same form as the percentage return on capital expenditure expected in commercial investment schemes. In the case of the commercial investment scheme a comparison of the expected rate of return with the current rate of interest is, of course, valid because it will reflect the likelihood of the undertaking being able to meet its financial obligations to those from whom it borrows capital resources. However, in the case of commercial investment, the benefits to the community from it will inevitably in real life be shared between the customer and the owner of the capital equipment.

The owner will, however, only bring to account in forecasting the return on capital the share of the benefits accruing to him. There will thus be benefits to the community in addition to the benefits brought to account in the commercial return.

Accordingly for two investment schemes of comparable value to the community, the return on capital shown in the case of one of them by a social benefit calculation will be larger than the return shown by a straightfoward commercial calculation for the other. In making use of the social benefit calculation as a guide to investment decisions (or as a *post hoc* justification) it is therefore important to know by how much the one will exceed the other.

The adoption of the social return criterion for road investment and the commercial return in other cases would have the result that the return on the marginal road investment would be lower, perhaps much lower, than the return (calculated on a comparable basis) on the marginal investment in the rest of the economy. On the other hand the adoption universally of the social return instead of the commercial return as a guide to investment decisions would add greatly to the volume of investment regarded as desirable unless at the same time different standards were to be adopted for the amount by which the expected return should exceed the rate of borrowing.

The report by the Road Research Laboratory on the London–Birmingham motorway contains the following remarks (page 63):

' In order for the construction of the motorway to be worthwhile, it must be shown that the rate of return obtainable is greater than the current rate of interest and (more rigorously) greater than the rates of return obtainable in other uses of capital (including other road improvements).'

If the arguments I have put forward are right, it is to be questioned whether the condition that the social rate of return should exceed the rate of interest is in any sense sufficient to establish the desirability of the investment. It is, of course, a necessary condition, in the sense that in order to compete in the Olympic High Jump event a competitor must be able to jump two feet.

The sufficient condition is, of course, much more rigorous. Investment schemes where the expected commercial return is less than twice the rate of interest will often be rejected. Such rejected schemes on a social benefit calculation will show a higher return, in many cases higher than the return on road investment schemes which, if it were to be assumed that the social return is in all respects comparable with the commercial return, would be regarded as desirable.

The Road Research Laboratory study, however, does not consider the price the user of the motorway is to be charged, nor does it deal with toll questions. In practice, of course, fuel duty paid in respect of the fuel consumed while on the motorway can be regarded as the price paid for its use. In the M1 case if the forecasts of traffic volume are achieved, the road user seems likely to pay a price which provides a reasonable return on capital employed, though the cost of the 70 miles so far built may in fact be about £30 million. This, of course, ignores the fact that for other parts of his journey the user may not be paying the economic price for the roads he uses. The volume of traffic on M.1 and the return by way of fuel duty would be lower if more were to be charged for other parts of the journey. In other cases, however, the fuel tax to be paid may be negligible in relation to the cost of the road asset which it is proposed to construct; for example, a bridge as a result of which journey times are greatly reduced may be very expensive in relation to its length and the fuel duty paid in consequence of its use will therefore be quite inadequate as a price for its use. There is, of course, no inconsistency, even with the present system of charging for road use, in levying a toll in such cases; the proposal to do so in the case of some future bridges and tunnels in this country is economically quite sound and there is, of course, little practical difficulty in doing so. In other cases, such as urban road improvements, the capital costs of the new works will be very high and the fuel duty to be paid by the potential users will be too low a price. There may be practical difficulties in instituting specific tolls or other economic charges in such cases but the alternative of not charging would appear to be more expensive in the end. Either capacity, which would have been adequate to meet the effective demand for road space if the economic price were paid by users, will be inadequate and congestion will be the result; or else heavy additional investment will need to be incurred to eliminate congestion, yielding much lower benefits to the community.

Inevitably, we are led to the conclusion that trying to cure road congestion simply by increasing road capacity is very much more expensive than the use of a combination of an economic price with the additional investment that effective demand at that price shows to be worthwhile. It can also be pointed out that a given investment expenditure will substantially reduce road congestion in many more places if a proper price is charged than if no attempt is made to recover from the particular users the cost of the improvement. (In deciding on the 'required' capacity of the road works, there is, of course, a need first to decide whether a toll is to be charged or not.)

However, the existence or otherwise of an appropriate price structure, though it is fundamental to rational investment decisions, does not alter the fact that the social return is a different quantity from the commercial one. This may be illustrated by two large investment projects which are of topical interest. The first is the addition to the London Transport railway network known as the Victoria Line. As a strictly commercial investment it is clearly one to be rejected. With the present charges for road use in London, the users of the Victoria Line by themselves would not be prepared to pay a price which would enable London Transport to recover the additional interest charges arising from the investment. On the other hand, those who have examined the project in relation of the London traffic problem have concluded that it would provide great benefits both to road users and to users of other underground railway services. The implication is that, on a social benefit calculation, taking into account factors such as the savings in time by those who transfer to it and the savings in time to those whose journeys elsewhere are less impeded as a result of such transfers and taking account of the benefits in enabling additional traffic to move, the investment in the line shows a net benefit to the community.

This, of course, raises in question the justification for judging transport investment, either in road or in rail, by the social criterion, if other investments equally socially desirable are to be judged by a commercial criterion and rejected.

The other example is the case of the Channel Tunnel project. Here we have a project which on a detailed assessment of its commercial prospects is expected by some on a commercial basis to yield, under fairly favourable economic conditions, a substantial return. No published estimate has to my knowledge been made of its social return on the motorway basis of calculation, but there is little doubt that this return would be substantially higher than the commercial return in view of the savings in time and other advantages over existing methods of crossing the channel. Its construction, however, raises a number of political questions and would require some investment financed from government sources and there are doubts whether it will be undertaken. It would, however, clearly be wrong to reject it on the basis that the expected commercial return is lower than the social return forecast for major road construction projects in this country, which would yield on a commercial basis a smaller return than a channel tunnel.

Finally reference may be made to a typical electrification scheme for a suburban railway service. For one scheme I have attempted a calculation of the social return, using the same factors as are used in the M.1 study to evaluate not only time savings but also the other social benefits. This I have compared with the commercial return calculated in the normal way. Depending on the view taken of likely traffic, the commercial return ranges between 7 and 10 per cent, but making the same assumptions about traffic the social return would be some 30 per cent. Varying the view taken about future events and about fare levels would naturally yield somewhat different figures. However, for investment schemes where passenger movement is largely concerned I think the social return will seldom be less than twice the commercial one. That there are benefits to the users over and above the amounts they pay in fares to the railways after electrification is suggested by the greatly increased property values which have followed the extension of railway electrification, of which a recent example is to be found in East Kent.

To sum up, there is no reason to reject the principle of the social return calculation as a method of assessing how worthwhile a road investment is. Sensible changes should first be made to the road tax structure and when the consequent effects on road vehicle use have become apparent, traffic studies and forecasts which assume that an appropriate price will be charged can be made. From these studies calculations of social benefits can then be made. Once this has been done, there remains the need to be aware how many times greater than the currently accepted commercial return the social return should be to make a road investment scheme as desirable a use of capital resources as investment projects currently undertaken on a commercial basis.

SURPLUS CRITERIA FOR INVESTMENT

By C. D. FOSTER

Since Dupuit developed the notions we know as consumers' and producers' surpluses from Adam Smith's distinction between value in use and exchangeable value,[1] there has been a close association between the discussion and use of social and consumers' surplus criteria for investment, and investment in roads and roadworks. Indeed they have never found much popularity elsewhere. In recent years they have often been discussed in connection with the theory of investment in roads.[2] And in practice very rough approximations to these criteria have been used to work out a rate of return on road investments.[3] The example which is of most interest to us in England is the Road Research Laboratory's calculation of a rate of return on the London-Birmingham Motorway. In the course of a description of its calculations the Laboratory indicates the ideal rate of return it would calculate if it could. ' The economic value of a proposed investment can be expressed as the present value of the future stream of *net benefits accruing to all affected by it*, divided by the present value of its costs over its life.[4] Less rigorously the value of the investment may be indicated by expressing *net annual benefits* as a rate of return on the capital cost of the motorway (or by estimating the pay-off period, i.e., the time taken for the future benefits of the motorway discounted back to the time of construction to reach a sum equal to the capital cost).'[5] This ideal can, I believe, be fairly described as the social surplus criterion for investment.[6] But it is only an ideal. As the Laboratory acknowledges, the actual rate of return it uses is a very rough approximation to this criterion.

[1] J. Dupuit, ' On the Measurement of the Utility of Public Works ', *Annales des Ponts et Chaussées*, 2nd series, Vol. 8, 1844, reprinted in *International Economic Papers*, No. 2, 1952.

[2] E.g., H. Hotelling, ' The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates ', *Econometrica* (1938); Beckman, McGuire and Winsten, *Studies in the Economics of Transportation* (1956); Meyer, Peck, Stenason and Zwick, *Economics of Competition in the Transportation Industries* (1959), ch. 4; and an article by A. A. Walters, ' The Consumers' Surplus Criterion for Investment in Roads ', to appear shortly, which the author has kindly shewn me before publication.

[3] E.g., ' The Economics of Highway Planning ', *Technical Bulletin No. 7*, Oregon State Highway Dept., (1937), which was one of the first; *Road-User Benefit Analysis for Highway Improvements*, American Association of State Highways Officials (1952) which sets out a calculus to be used by Highways Officials to help them work out a rate of return; W. D. Ross and L. J. Melton, *Financing Highways Improvements in Louisiana* (1955); H. W. Bevis, ' The Application of Benefit-Cost Ratios to an Expressway System ', in *Proceedings of the Highway Research Board* (1956); *London-Birmingham Motorway: Traffic and Economics*, Road Research Technical Paper No. 46, D.S.I.R., H.M.S.O. (1960).

[4] Thus expressed the criterion is incalculable since no limit is put to the period over which net benefits are to be summed. We need to assume a cut-off point beyond which the stream of net benefits will be disregarded to make the economic value of an investment finite. Cf. J. de V. Graaff, *Theoretical Welfare Economics* (1957), pp. 93–6.

[5] *Op. cit.*, p. 43. My Italics.

[6] This term, social surplus criterion, is borrowed from Professor Hicks, who used ' social surplus ' in this sense: ' The social surplus = consumers ' surplus + producers' surplus—loss of potential surplus on other commodities ' where all surpluses are classified as consumers' or producers' surpluses. J. R. Hicks, ' Rehabilitation of Consumers' Surplus ', *Review of Economic Studies*, Vol. 8, p. 116.

It is, in fact, a rough approximation to a consumers' surplus criterion, modified even so by many simplifying assumptions.[1]

The Road Research Laboratory goes on to claim that its ideal rate of return is comparable with other rates of return. ' The rate of return may . . . be used for comparison with other forms of investment and with the rate of interest to assess whether the investment (in a motorway) is worthwhile.'[2] Mr. Glassborow has challenged the legitimacy of comparing a rate of return of this kind with what we ordinarily think of as a rate of return on capital, in another article in this symposium.[3] My purpose here is not dissimilar from Mr. Glassborow's. It is an attempt to clarify this question of comparability of rates of return.

But it must be admitted that at present the issues discussed in the rest of this paper are academic. Perhaps the most significant statement in the Road Research Laboratory's monograph is this one: ' The study was not undertaken with a view to deciding whether or not the London-Birmingham Motorway should be selected as a profitable scheme, but it was undertaken as a subject of research to see whether reliable methods of assessing both the traffic that would flow on it and the economic value of the scheme could be devised.'[4] Not only does it appear that the decision to build the London-Birmingham Motorway was not an economic one, but that subsequent decisions to build major roads have also not been based on economic assessment. The costs of building roads and road improvements are in most cases easily discovered and well known. But the benefits which the Laboratory considered in working out a rate of return for the London-Birmingham Motorway require laborious and costly research.[1] Therefore although we can perhaps be sure that all roads built so far or projected would show a positive rate of return of the Road Research Laboratory's type, if the necessary research had been done to make the calculation possible, there is no evidence that any rate of return has been worked out for them. And it is almost certain that there are other projects than those undertaken and intended which would show a higher rate of return, calculated on the same basis. Most of these would be less spectacular—road-widening, the improvement of intersections, etc.—or the building of urban motorways. In practice, the economic planning of the roads programme seems to be more primitive than the economic planning, for example, of British Railways or the National Coal Board. Almost all important decisions are taken by the Ministry of Transport, by itself or in consultation with local authorities. We have much less evidence on the criteria used to decide investment priorities for roads than we have for the nationalised industries, because the report of the Select Committee on Estimates which investigated the building and planning of trunk roads was superficial where it touched on economic issues,[2] and did not yield anything like the wealth of information to be found in the reports of the Select Committee on the Nationalised Industries about the National Coal Board and British Railways. So we are dependent mainly on press reports and other, various, mostly fragmentary, information for what we know of the criteria by which decisions are made. It seems that in most cases the criteria governing the size, quality, and therefore the cost, of new roads and road improvements are technical; that the choice of projects is based on a rough estimation of traffic flows and what the public needs; and there is interplay between the Ministry of Transport and local authorities which has meant that an enterprising and persistent local authority can get a project undertaken sooner than would otherwise be the case.[3] Although we may hope that the time will come when the priorities of the roads programme are decided on better evidence, at present it would be foolish to delay road-building until the machinery for making better calculations exists. This was recognised by the Select Committee on Estimates in 1959.[4] But one can only hope that the Road Research Laboratory's resources will be increased so that it can undertake this kind of investigation for all the more important projects, but, of course, before, not after, decisions are made. The Select Committee recommended this. It does not seem to have persuaded the Minister of Transport that it is urgent, which it surely is. Until this is done, there cannot be any programme worthy of being called a Roads Plan, and *a fortiori* a Transport Plan. (Neither can we decide sensibly how large British Railways should be and what modernisation is worthwhile until we have such a Roads Plan.) In the rest of the paper we are looking forward to a time when the Roads and overall Transport Plan of this country is more carefully worked on the lines indicated by the Laboratory's calculations.

---

[1] This is also the case when the calculation is expressed not as a rate of return but, as is more usual in the United States, a benefit-cost ratio. The ratio of net benefits expected to accrue from a highway to the costs of its construction is used to rank investment opportunities in an order of priority. Here again we get the distinction between the ideal and the actual. For example, Bevis classifies benefits in four groups: 1. Direct-user; 2. Indirect-user; 3. Direct non-user; and 4. Indirect non-user. ' Direct-user ' benefits are those which accrue to users of the highway and can be measured in direct money terms. Examples include reductions in fuel and oil consumption, maintenance costs, losses due to accidents, and wages to operators of commercial vehicles. Indirect-user benefits include those intangibles accruing to users for which some money value must be assigned. For example, they may be reductions in driving time, decreased driver strain, increased comfort, and so on. Direct non-user benefits are those received by people who do not use the facility but whose livelihood is improved by it . . . Indirect non-user benefits are those intangibles accruing to the community through improved transportation facilities. For example, juvenile delinquency and the crime rate may fall as blighted areas are given a new lease of life. Or the ability to evacuate a city faster in the event of atomic attack has some potential value in the potential saving of human lives.' Bevis comments that ' the measurement of all of these benefits constitutes an enormous task. Such a computation is probably prohibitive from the standpoint of current technology '. Social costs should also be allowed for. Therefore ' social or non-user benefits should be excluded from consideration until such time as the technology is sufficiently advanced to provide precise measures for both these types of benefits and for social costs '. (*Op cit.*, pp. 65–6).) Once more the actual ratio is used a rough approximation to what an economist would call a consumers' surplus criterion. Cf. A.A.S.H.O., *op. cit.*, pp. 2, 6–16, 20–2.)

[2] *Op. cit.*, p. 43.

[3] ' The Roads as Economic Assets: Investment Decisions ', *supra.*

[4] *Op. cit.*, p. iii. Also p. 43.

---

[1] These are: (i) value of savings in persons' working time; (ii) in vehicle time; (iii) in fuel consumption; (iv) in tyre-wear; (v) in brake-wear; (vi) in clutch-wear; (vii) in the cost of accidents; (viii) in non-business time; both for traffic already travelling on re-existing roads for which it pays to divert to the motorway and for new, or generated, traffic. *Op. cit.* See also *Assessment of Priority for Road Improvements*, Road Research Paper No. 48, D.S.I.R. H.M.S.O. (1960).

[2] These criteria are mentioned particularly in the answers to questions 246 and 247 in the Minutes of Evidence. They are left vague. *First Report of the Select Committee on Estimates: Sub-Committee F* (June 10th, 1959): Trunk Roads.

[3] *Op. cit.* Report para 12. The best-known case was the Preston by-pass.

[4] *Ib.* Report para 7.

This problem of comparability of rates of return is not confined to comparisons between so-called ordinary or private rates of return and the surplus criteria commonly used on road investment. For example, it often seems to be argued or implied that we can compare a private rate of return with the rate of return on an investment on a nationalised industry or public corporation. This can be so only when there is no element of subsidy in the pricing, investment and purchasing policies of a nationalised industry. Formally the fact of comparability in such circumstances is unexceptionable, given that the aim of all investors is to maximise the rate of return on investment. Comparability is established by definition.

But to establish comparability in fact we must consider what it means to ' get rid of all elements of subsidy '. To get comparability it is necessary to define a subsidy as follows: any revenue which it is within the power of a firm (or nationalised industry) to earn, but which it chooses not to demand; any costs which it is within its power not to incur but which it chooses to incur. By this definition a subsidy covers more than is often meant. Obvious subsidies are the surplus of cost over revenue of running an unremunerative branch line; that which is given to a manufacturer when a firm buys from him when it could buy the same good from another more cheaply. Subsidies which are possibly less obvious are those a consumer gains by being charged a standard rate when he would be prepared to pay more and a firm could extract this extra from him by discriminating between him and other consumers. Taking the Central Electricity Authority as an example of a nationalised industry making a profit, it is in this sense a subsidy to consumers to use a profit surplus to lower prices when they would pay a higher price which would yield a higher profit. Nationalised industries when they make profits are ordinarily expected to pass on some of the benefit to the consumer in this way. In so far as they do, their rates of profit are not comparable with other rates of return.

For example, let us suppose that British Railways has been revived by closing lines and modernisation so that it makes a profit after it has paid its capital charges. If this were the case, many would argue that prosperity should be passed on to the consumer rather than go as a dividend to the Treasury. Let us suppose that this advice prevails and that the average rate per passenger mile is reduced from 3d. to 2d. By intention this is less than the railways could get for their services. In other words they are paying a subsidy of a penny a mile. In this situation, it would be ridiculous to compare the rate of return on the railways' capital with that earned by a private firm which did try to maximise its profits.

Now let us suppose that the question is asked if some station should be modernised and a rate of return is worked out on the investment. Part of the economic advantage is to come from lower costs of operating the station. This will yield a 4 per cent rate of return on capital. Another part will come from new passenger business. The yield arising from that is expected to be 5 per cent—9 per cent in all. If we were then to compare the rate of return on this investment with one of equal risk contemplated by I.C.I. on which the return was expected to be 10 per cent we would be comparing the incomparable. It might be the case that the rate of return would go up, say, to nearly 12 per cent, if the passenger rate per mile were increased to 3d., which we will take to be the rate which would maximise profits. Then we would be comparing like with like, if all elements of subsidy were absent from each rate of return. But to compare the 9 per cent and the 10 per cent is plainly absurd, since British Railways by assumption has chosen to charge less than it could: and that is why in this example its rate of return is lower than I.C.I.'s.

We can go further than this and say that it would be self-contradictory for a nationalised industry which charged less than the traffic could bear to take as its criterion for investment the commercial private or profit-maximising rate of return. When a firm sets out to maximise profits this is equivalent to its saying that the only beneficiary of the investment in which it is interested is itself as entrepreneur or profit-earner. But if it imposed on itself an ordinance of self-denial and, for whatever reason, took account of the interest of consumers or others when fixing its prices it too would be behaving illogically if in deciding on its investment it considered only its own profits.

Therefore it would make sense to compare rates of return in private and nationalised industries only if nationalised industries had the same policy as private enterprise, that is, generally, profit maximisation. While nationalised industries are in deficit many would argue that they should charge what the market can bear and take every measure to minimise losses and make a profit. But few would wish to go as far as insisting that nationalised industries *in surplus* should maximise profits, however much they desire ' commercialisation '. And it makes sense to compare rates of return to different nationalised industries only if they are pursuing the same policy. It makes no sense, for example, to compare the rate of return in the electricity industry with that in British Railways, since the former, being in surplus, has a policy of passing on some of its profits in lower prices.

It might be objected: how can we say that private or profit-maximising rates of return are comparable when the degree of monopoly is greater in some cases than in others? This cannot be an objection to the notion of comparability as it confronts private firms or investors, since their object is to maximise their profits given the conditions (constraints) under which they operate. But it may seem odd to say that the rates of return of a private firm and a nationalised industry are comparable only when they both have the same pricing and investment policies—e.g. when they both try to maximise their profits. This might seem to suggest that the rates of return on private and public enterprises would be comparable only if public monopolies were allowed to earn the *large* monopoly profits their presumed monopoly powers would enable them to earn.[1] This would be a misunderstanding of the point at issue. Legislation—anti-monopoly, public health, factory, etc.—is one of the factors affecting firms' ability to earn monopoly profits. Therefore if legislation permits differences in private rates of return the presumption is (i) that the permitted differences are socially acceptable (if not desirable); or (ii) though not judged acceptable, the government is powerless to legislate against them. In other words the purpose of legislation is to fix limits (constraints) within which firms are free to maximise their profits or pursue any other policy they choose. It establishes the conditions of comparability. If the government should feel that, given present legal constraints, the private rate of return would be too high in a nationalised industry, either for reasons of resource allocation or on other grounds, it would impose constraints—formal or informal—on the ability of the nationalised industry to maximise profits, just as it could make a similar decision in respect of a private firm. But this does not alter the fact that it is self-contradictory for a firm to impose constraints on itself to serve a policy other than profit maximisation and then use a profit-maximising rate of return to select investment opportunities. Nor does it make it not self-contradictory for the state to require a public enterprise to have a policy other than profit maximisation and then to judge the desirability of investment in it by a profit-maximising rate of return. If the purpose of the regulations and pressures affecting the pricing and investment policies of public enterprises is to set limits within which they are to maximise their profits, then the state has established the conditions of comparability. The rates of return on private and public investment are comparable. (If this is the case, a requirement that a public enterprise should break even, that is, not make a profit, is tantamount to the social judgment that there should be no more investment in that enterprise.) But if the purpose of these regulations and pressures is that public enterprises should pursue a policy other than profit maximisation—e.g. that they should be run in the consumers' interest—then their private or profit-maximising rates of return are not comparable with those of profit-maximising firms. The second of these alternative purposes is surely nearer the truth.

We now turn to another aspect of the problem. This is where different enterprises—public or private—are pursuing different policies and are not using the same rate of return. We will consider two questions: (i) the rationale which might be expected to underlie the use of investment criteria other than profit maximisation; and (ii) if there is anything we can say about the principles on which resources can be allocated between enterprises using different types of rates of return. The second is Mr. Glassborow's problem.

---

[1] Although this has nothing to do with the theoretical point at issue, there is some reason to believe that the term ' public monopoly ' is an archaism. There are some nationalised industries which have potentially considerable monopoly power: e.g. the G.P.O. as post office and as provider of telephone services, I.T.A., and the Ministry of Transport as provider of highways. But the more important nationalised industries are monopolies now only in name. British Railways competes with road services, airlines and canals; coal competes with electricity, gas and oil for most purposes; overseas airlines compete with foreign airlines to some extent.

There is a vast family of criteria which can be used to choose between investment opportunities, of which profit maximisation is the most familiar. They can all be called surplus criteria. We will only consider two besides profit maxisation: the consumers' surplus and the social surplus criterion. These are the most interesting, both in general and for road economics.

What is meant by ' surplus ' and therefore by ' surplus criterion'? Let us begin by recalling some well-established propositions about surpluses. Consumers' surplus, to begin with that concept, is one species of economic rent, as Marshall was at pains to point out.[1] There are also factor-owners' surpluses, workers' surpluses, lenders' surpluses among others. There are surpluses wherever a rent is earned. There is also one kind of surplus ruled out by definition under perfect competition but which has widespread existence in reality: producers' surplus or monopoly profits. It is this surplus which is relevant to the profit maximisation criterion for investment. Using factor in the widest sense, a surplus, or rent, is gained by every factor not in perfectly elastic supply. Similarly, a consumer's surplus is gained by every consumer whose demand curves are not perfectly elastic[2] (provided he is not faced by discriminating monopolists of what Pigou called the first degree, who are able to discriminate not only between consumers but between units sold to consumers, so expropriating all his consumer's surplus).[3] Just as factor-rent, or surplus, does not enter into costs of production for the owner of the factor, so consumer's surplus, or rent, does not enter into the costs of his consumption. Factor-rent is the difference between the payment a factor of production receives and that beneath which it would cease to be employed. By analogy, a consumer's surplus is the difference between the income he actually receives and that beneath which he would cease to live: subsistence-level income. The consumer's surplus on a commodity is the difference between the price paid for it and the price at which a consumer would substitute something else for it.

There is then a large number of possible surplus criteria: (i) there are as many criteria as there are factors of production which have surpluses to be maximised; (ii) there are also those criteria which would maximise the surpluses of more than one factor of production; (iii) there are as many consumers' surplus criteria as there are commodities multiplied by the number of consumers of each; (iv) again there are also those which would maximise more than a single consumer's surplus; and lastly, (v) there are those criteria which would maximise some combination of producers' and consumers' surpluses. (And this describes the range of surplus criteria only loosely because of the different weights which can be given to surpluses. This will be considered further below.)

---

[1] *Principles of Economics*, Appendix K.

[2] Cf. H. Myint, *Theories of Welfare Economics* (1948), p. 148.

[3] *Economics of Welfare*, pp. 275–9. Similarly discriminating monopsonists can expropriate other surpluses.

By the social surplus criterion I mean one which chooses between alternative investments by comparing what are expected to be their effects on all surpluses, in a society, by some definition of society, usually a nation. All surplus criteria which have been used, or could be used, fall short of this: (i) occasionally from a deliberate value-judgment but more often (ii) for practical reasons. (i) The effect of an investment on some people's surpluses may be ignored or treated differently from its effect on others' because of a value judgment on the desirability or relative desirability of some and the undesirability or relative undesirability of others. (ii) Though it has been shewn that it is theoretically possible to calculate the effect of an investment on all surpluses, in practice this cannot be done.[1]

But before considering this further we will explore *a* consumers' surplus criterion. It is not *the* (total) consumers' surplus criterion usually considered by economists which is a much more complicated concept. It is however the consumers' surplus criterion which appears most obviously relevant to the investment and pricing policies of nationalised industries and the roads programme.

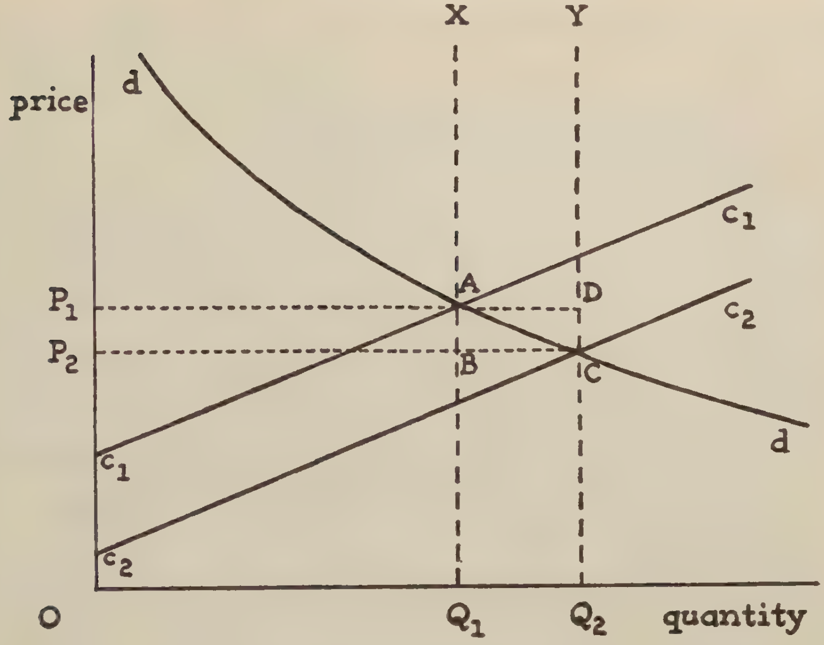
We have already considered situations in which it appears self-contradictory for a firm, or nationalised industry, to use a profit-maximising rate of return while at the same time placing a self-imposed constraint upon its ability to maximise profits. What rate of return should such a firm or industry use? We will make the supposition that runs through much past thinking on the function of nationalised industries, which is that they are run for their consumers' benefit.[2] Let the policy be that all profit should be paid out in reduced prices so that the industries make no profits. To simplify the discussion we will assume we are talking of one nationalised industry producing one product whose profits are taken out by uniform reductions in its price. It might be thought that such an industry, planning to make no profits, had no way of choosing between alternative investment opportunities. But that would be unimaginative. Since the aim of its existence is to benefit the consumer by price reduction, the logical implication of this policy would appear to be that it should distinguish between investment opportunities by the benefit they are expected to confer on consumers. The benefit expected from one of these investments is illustrated in the diagram.

The effect of making this investment is to reduce the cost of producing the nationalised industry's product from c_1c_1 to c_2c_2. When the costs were at their pre-investment level, consumers bought quantity Q_1 of the product at price P_1. They can now buy the same quantity at price P_2: so that the first part of the gain in consumers' surplus is measured by the rectangle

---

[1] H. Hotelling, 'The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates', *Econometrica* (1938); and among Professor Hicks' writings on the subject, particularly chapters 10 and 18 of his *Revision of Demand Theory* (1958).

[2] Cf. John Hughes, *Nationalised Industries in the Mixed Economy*, Fabian Tract 328 (1960), who considers in detail what this has meant for the nationalised industries.



P_1ABP_2. The second part of the gain is much harder to estimate. Because of the lowered price, old consumers buy more of the product and new consumers begin to buy some of it. The post-investment quantity demanded in Q_2. But we cannot measure the benefit to the consumer by multiplying P_1P_2 by Q_1 Q_2. It is in fact the area under the demand curve, dd, taking any path within the rectangle ABCD. It is plausible to imagine the demand curve, dd, taking any path within the rectangle ABCD between A and C which did not interest itself. (It is logically possible it should have a path outside ABCD but between XQ_1 and YQ_2 but this is most unlikely and normally the possibility can be ignored.) Therefore it is necessary to know the shape of the demand curve between A and C to calculate the second part of the gain in consumers' surplus accurately. This is information which one cannot expect a firm or nationalised industry to have and it is information which it would be difficult and dear to get. But unless there are special reasons for thinking that the demand curve has a markedly different shape, the best thing a firm or nationalised industry could do would be to make the rough assumption that AC was a straight line and calculate the second part of the gain in consumers' surplus as ABCD/2. The gain in consumers' surplus is therefore:

$$P_1ABP_2 + \frac{ABCD}{2}.$$

It we then suppose that the nationalised industry has gone through a similar calculation in respect of another investment opportunity and has produced a smaller gain in consumers' surplus and so a lower rate of return on the cost of the investment. It would follow that by this criterion, the consumers' surplus criterion, the first investment is preferable to the second. It also follows that a nationalised industry whose policy was to pass on benefits to its customers in this way would be contradicting that policy if it did not select the first investment opportunity. (If the demand curve had the shape ABC between A and C, this would subtract something from the rate of return on the first investment and bring its yield below that on the second investment. Conversely the rough assumption might under estimate the second part of the consumers' gain on the second investment. The moral is that the closer together the rates of return are the more probable it is that such a mistake might be made. But this is of course a risk the investor takes, if he makes the rough assumption.)

This use of a consumers' surplus criterion rests on certain assumptions about what it is desirable to consider in estimating a rate of return.

(1) It rests on the assumption that the only benefits to be considered are those of consumers.

(2) It assumes that a pound's benefit accruing to any one consumer is to have equal weight in estimating the rate of return as a pound accruing to any other.

(3) It assumes that any pound, let us say the *n*th pound, of benefit accruing to any consumer is to have the same weight as any other pound of benefit accruing to him.

(4) It assumes that a pound of benefit accruing to someone at any time, *t*, is to have the same weight in the calculation of the rate of return as a pound accruing to him, or anyone else, at any other time.

The sensible attitude towards these assumptions is not to criticise them because they are false. They cannot be either true or false, since they are not statements of fact but value judgments. The second assumption does not assert that a pound to one man gives as much benefit to him as it does to anyone else. What it means is that any differences in the benefit or satisfactions different people derived from a pound are to be neglected. It would be possible in principle though perhaps difficult in practice to weight pounds of satisfaction accruing to certain kinds of consumer, say the poor, more heavily than pounds of benefit accruing to other consumers. Similarly, the third assumption does not mean that the utility of income to a man is constant; but that in calculating a rate of return the decision is made to neglect and inconstancy in the utility of money. The fourth assumption does not mean that the rate of time preference of consumers is zero; but again that this assumption is made in calculating the rate of return, either because of a deliberate value judgment or because it is impracticable to make allowance for another rate of time preference. These assumptions are normative not descriptive. Of course similar assumptions to the first, third and fourth of these are needed to support the ordinary calculation of a private or profit maximising rate of return.

There are two points that should be made here:

(1) We have ruled out another possible source of self-contradiction by the assumption that the price charged to consumers covers costs, that is long-run average costs, including depreciation on a replacement cost basis, the cost of finance, and a premium for risk. If the price were to be lowered below this, then the possibility of a self-contradiction could arise. Part or even all the cost of producing the product would have to be borne by others than the consumers. Let us assume that the difference between total revenue and total cost is paid out of taxation. We will assume (i) that there are taxpayers who are not consumers of the product; (ii) that tax paid by any taxpayer in respect of the cost of the product does not equal the cost of the quantity of the product he consumes. The effect of such a policy is to redistribute income from taxpayers as a class to consumers of this product as a class. But in so far as taxpayers are consumers the benefit they gain from a fall in its price is offset by the increased amount they pay in taxation to meet its cost; so that the gain in their surplus is lower than it otherwise would be. If it were expected that this would be important, as indeed it might be with a large roads programme, allowance should be made for it in estimating a consumers' surplus rate of return lest there should be a policy self-contradiction. Again this possibility is not peculiar to the consumers' surplus criterion. It could happen with the profit maximisation criterion, in so far as those who receive the profits of an enterprise are consumers of the product; but of course the benefit to them is more likely to be increased by a lowering of the price of the product as a result of the investment than diminished by an increase; and in most cases the difference will be insignificant.

(2) The second point is that the abandonment of a pricing policy where price is fixed at or above marginal cost may mean that there will have to be rationing. As an example let us take the limiting case and assume that the price of the product is set at zero. Therefore the marginal cost to the consumer of procuring a marginal unit of the product is zero; but the marginal cost of producing a marginal unit is above zero. Now if the consumers' surplus criterion is used, output demanded is determined not by the intersection of the marginal revenue curve and the producer's marginal cost curve, but by the intersection of the demand curve and consumers' marginal cost curve, which since their marginal costs are zero is the same as their average cost curve.

There will be a limit to the quantity demanded only if the demand for the product at zero price is not infinite. That there should be a point where the marginal valuation of the product should fall to zero and therefore equal the consumers' marginal cost is a reasonable assumption for all products. So we may assume that there is some limit to the quantity demanded. Given that there is a limit to the quantity available there are three possibilities: (i) the output available of the product is less than this limit so that, at least in the short run, the product will have to be rationed; (ii) it is possible to produce the output desired by reallocating resources from other sectors of the economy where there is no rationing problem, i.e. sectors in which a surplus criterion with price less than marginal cost is not followed; (iii) the sectors of the economy in which the rule is followed are too large for the required output in each to be met by reallocation of resources; so that there is an overall shortage which will have to be met by rationing if not for the product in question, then for other products.

The first point is not relevant if the price of the product is equal to or above producer's average cost. The second point is not relevant if the price of the product is equal to or above producer's marginal cost. It is therefore possible to use a (modified) consumers' surplus criterion without running into either of these difficulties. It is a modified criterion as the rule is that consumers' surplus is to be the sole measure of benefit for an investment only in so far as price is equal to or above marginal cost if marginal cost is above average cost and equal to or above average cost if average cost is above marginal cost. But the adoption of the modified rule is not an absolute good. We may prefer not to modify the rule and face the difficulties.

So far we have considered two bases for the calculation of rates of return: profit maximisation and consumers' surplus. However these are only two members of the vast family of surplus criteria. It is sometimes argued that they are incomplete criteria. Such an argument presupposes that there is a complete or ideal criterion: which is usually said to be the social surplus criterion.

For instance it is often objected to both the profit maximisation and the consumers' surplus criteria that they leave out many categories of social cost and social benefit and are therefore incomplete. One example is that neither private firm nor public enterprise take into account any loss of revenue they may inflict on competitors when they make an investment. A recent example of this is the effect of the opening of the London-Birmingham Motorway on the transport cafés of Towcester. By diminishing the volume of traffic on A5, it has taken away some of their business and has forced some to close down, with consequent multiplier effects upon the prosperity of the town of Towcester, which depended, apparently, to some considerable extent upon these cafés. This is a cost, social cost, of the London-Birmingham Motorway which was omitted in working out the motorway's rate of return. And it would not be hard to find many similar effects of the investment of private firms and public enterprise, intended and unintended. It is often argued that such social costs (or conversely benefits) *should* be taken into account. Sometimes it is said that both private and public enterprise *should* allow for them. More often it is said that although one cannot expect private firms to do so, nationalised industries should because they are agents of society and should take into account, in so far as is practicable, all social benefits and costs.

If we accept such arguments we are led to the social surplus criterion for investment, as an ideal. It cannot be achieved in practice because its calculation would require perfect knowledge of all demand and cost curves for all society's goods and services. But even if it were practicable to use this criterion, there is no compelling reason why we should accept it as an ideal.

Let us imagine that the ideal is practicable. We assume perfect knowledge of every demand and cost curve. We calculate the effect that an investment will have on all surpluses. We consider not only the consumers and the entrepreneurs but everyone who is affected by the investment whether as manufacturer of a competing, or complementary product, as owners of factors of production for the firm contemplating the investment, or any other firm affected by the investment; e.g., it will lower producers' surplus in respect of close substitutes; it will raise it in respect of complementary goods; it will have various effects on workers' surpluses, consumers' surpluses, savers' surplus, etc., throughout the economy. On the assumption of perfect knowledge all these effects have calculable magnitudes. We then perform the addition and subtraction of the Hicks criterion:

'The social surplus = consumers' surplus + producers' surplus − loss of potential surplus on other commodities.'

If the result of our calculations is positive we may conclude that it is beneficial to invest. The amount by which it is positive will give us the social surplus rate of return on the investment. If we had a number of investments which we wished to arrange in a schedule of priority the social surplus criterion would do it for us.

We do not have to accept the social surplus criterion rate of return as an ideal because it rests, like the consumers' surplus criterion, on certain assumptions or value-judgments. The value-judgments are similar:

(1) We assume that every surplus is to be considered in judging the desirability of an investment.

(2) We assume that any man's pound of benefit is to have equal weight with any other man's pound of benefit; and similarly that any man's pound of cost is to weigh the same as any other man's.

But on reflection these are curious assumptions because they take no account of income distribution. Let us imagine two investments. We have calculated the social product or benefit, and the social cost of each according to the social surplus criterion. We will suppose that in both cases the social net product is the same; but in the one case all the benefits go to the one man who is the sole gainer while all the costs are borne by the poor, while in the other case all the benefits are divided among the poor, while the costs are borne by the rich. The social surplus criterion regards the worth of each investment as the same taking no account of what may be thought of the very different income distributions involved. If the first investment was expected to yield a social rate of return somewhat above the social rate of return on the other, it would declare in favour of the former. But we can imagine many instances as dramatic and less dramatic when society through its institutions would not accept the dictates of the criterion.

Moreover we can imagine choices where one investment might have a much lower social rate of return than another and be preferred because the income distribution effected by it was preferred.

The point is that there is no irrefutable reason why we should accept that a pound's benefit (or cost) is to count the same to whomsoever it accrues. Society—if society is the decision maker—can employ any system of weights it pleases. For instance, to take a trivial example, it could state that any fall in bookmakers' profits is not to be counted a social cost of a given investment but an equivalent benefit; this is an intelligible value-judgment. We could therefore imagine many value-judgments which would decide what surpluses and costs are to be considered in working out the (or a) social rate of return. One might be that only the effect on consumers and wage-earners (consumers' and wage-earners' surpluses) should be counted; another that all benefits and costs should be considered except any costs falling upon entrepreneurs. (That would be an implication of a policy designed to establish perfect competition.) Another might be that a fall in certain benefits might be counted a gain—once more it might be producers' or entrepreneurs' surpluses, monopoly profits, etc. An attractively simple system of weighting would be that a pound's loss in one of these surpluses should count as much in favour of an investment as a pound's gain, say, in consumers' surplus. We could of course imagine more complicated surplus criteria making finer discriminations between kinds of people. In all cases the essential distinction is twofold. (i) The first is between ' good ', ' bad ' and ' indifferent ' effects of an investment.[1] We can generalise this ' social ' surplus criterion:

The chosen surplus = (f_1(increase in ' good ' benefits)—
f_2(decrease in ' good ' benefits))—
(f_3(increase in ' bad ' benefits)—
f_4(decrease in ' bad ' benefits))

In order to make sense of any ' social ' surplus criterion or ' social ' rate of return we need a social decision which tells us what changes (plus or minus) are judged beneficial and which are not. (ii) A second set of distinctions is needed to tell us how to weight various benefits and costs in calculating the rate of return.

We can make similar criticisms of the supposed ' ideal ' nature of the other assumptions needed to underpin the social surplus criterion. These are:

(3) we assume that any pound's benefit, let us say the nth pound, accruing to anyone is to have equal weight with any pound accruing to him.

(4) we assume that a pound of benefit (or cost) accruing to anyone at any time, t, is to have the same weight in the calculation of the rate of return as a pound accruing to him, or anyone else, at any other time.[2]

---

[1] Cf. Graaff's statement that even at the ' engineering level ' value judgments are needed to classify goods into inputs and outputs. *Theoretical Welfare Economics*, pp. 15–6.

[2] Doubts about these cardinal properties of the Marshallian consumers' surplus criterion led Hotelling and Hicks to make their critical *ordinal*. For example Hotelling argues as

If we no longer think there is such a thing as an ideal social investment criterion there is no point in trying to include every social benefit and cost which it is practicable to estimate and give as our excuse that we are striving after the ideal. This matter of choosing a social rate of return has been handled in a more sensible way in connection with the problem of choosing investment criteria for underdeveloped countries;[1] though the distinction is not always made between what is omitted or included as the result of a value judgment and what is omitted or included on the grounds of practicability. Nevertheless these surplus investment criteria can make sense. They concentrate on the variables held most important for an underdeveloped country; first, a value for the private rate of return on an investment; secondly, a value for the effect on the balance of payments; others possibly for the effect on the level of employment; weighted by the selection of a rate of time preference; given a determinate value by the selection of a cut-off period which sets a limit to the time over which benefits will be counted. (It is usual to assume that money has the same value for everyone and constancy in the utility of money.[2])

It follows from our argument that no *a priori* reason can be given why a public enterprise should not use any surplus criterion the state pleases provided it is consistent in its use. (If the state wants a public or private enterprise to use some criterion other than a profit maximising one, and this should mean that it was unable to attract capital on the free market either at all or enough of it to finance the rate of investment the state thinks desirable, it is the state's problem to subsidise the cost of capital to it. Some rough notion of this sort presumably justifies the fact that public enterprises are permitted to borrow at rates of interest which represent a government subsidy to them. But it is rough justice, that is, no justice.)

It is proper for us now we have considered why there is no reason to hold that there is a right rate of return in any absolute sense, to examine some positive suggestions. Is there a case for rates of return diverging from the private or profit-maximising rate of return? The answer most people would give would change according as it were assumed (i) that all rates of return were

---

follows (w stands for social surplus): 'The change in w that will result from a proposed new public enterprise, such as building a bridge, may fairly be set against the cost of the bridge to decide whether the enterprise should be undertaken. It is certainly a better criterion of social value than the aggregate Σpq of tolls that can be collected on various classes of traffic. Determination whether to build the bridge by calculation of the revenue Σ pq is always too conservative a criterion. . . . *For society it is beneficial to build the bridge if w exceeds the interest, amortisation and maintenance costs.*' It is clear from our previous argument that we do not have to accept the criterion put forward in Hotelling's last sentence. But there is another point. The criterion seems to answer the wrong question. We usually do not want to ask whether an investment is ' beneficial ' or not—whatever our criterion, but if it is more ' beneficial ' than other possible investments; and it is thus in the nature of the problem that we are driven to require a more determinate method of distinguishing between investments, generally speaking, numerically determinate, that is, a rate of return.

[1] By A. K. Sen, for example, in his ' Some Notes on the Choice of Capital Intensity in Development Planning ', *Quarterly Journal of Economics* (1957); and his subsequent 1959 *Quarterly Journal of Economics* article; also, e.g., A. E. Kahn, ' Investment Criteria in Development Programmes ', *Quarterly Journal of Economics* (1951); H. B. Chenery, 'Application of Investment Criteria ', *Quarterly Journal of Economics* (1953).

[2] Sen, *op. cit.* (1957), p. 569, n. 1.

amenable to social policy in this respect; or (ii) that only the rates of return on nationalised industries, public corporations' investments and other public enterprises could be affected. The first assumption would require that private firms were taxed and subsidised or ordered to use rates of return other than profit maximising ones. This is not practical politics. Let us make the second. I would not myself argue that the macro-economic investment criteria which have been suggested by economists thinking of underdeveloped countries give us much useful guidance, for two reasons. (1) What matters to them matters much less in this country. There would not be much advantage to set off against the cost and difficulty of calculating the effect of investments on such variables as the level of (a) economic activity, (b) employment, (c) savings, and (d) the balance of payments. This would be true even if we made the first assumption. (2) It is arguable that it would be comparatively ineffective and wrong-headed to adjust the rates of return of nationalised industries and other public enterprises to such ends. It would be comparatively ineffectual because they have relatively little effect upon these variables in relation to their size. For example, the only important present problems of unemployment are local and the types of investment made by nationalised industries are rarely of much importance in eliminating these. Public enterprises do not in general have large propensities to import and export. It would be wrong-headed, I think, because any benefits are likely to be outweighed by distortion in the allocation of resources. I do not mean this in relation to any optimal allocation of resources, because we are not necessarily considering an economy near an optimum; but quite simply, altering the rates of return of public enterprises for these reasons would be likely to lead to a less than desirable rate of investment in the public sector. In principle it would be possible to compensate the public sector's rates of return so as to offset this, by weighting. I am not sure how one would set about establishing the principles of such compensation so that the desired rate (and allocation) of investment is achieved in the public sector. But it is surely more likely that any attempt to influence the rates of return of nationalised industries for these purposes would not be compensated in this sense but would accentuate the present unfortunate position where the investment programmes of the public sector are altered on grounds of national economic policy and this does lead to a less than socially desirable rate of investment there. (The burden the nationalised industries have suffered from in the past of having their investment funds cut or 'phased' when there has been a balance of payments crisis or inflation is in a very rough sense an adjustment to the criterion used for their investment.)

But if we are not going to bother about the effect of investments on these macro-economic variables, what divergences from a profit-maximising rate of return might we bother about? Here we meet the argument that public enterprises ought to adopt something approaching a consumers' surplus criterion on the grounds that they should be run in the consumers' interest. If this criterion is not to become unmanageable from impracticability it would be wise to circumscribe it carefully. It might be convenient (i) that it should be the direct consumer's surplus criterion only considering the consumers of the industry's products; (ii) that it should conform to the three assumptions about consumers' surplus given above so as to eliminate those problems of weighting (and incidentally be similar to a private rate of return in these respects); and (iii) that it should be the ' modified ' criterion given above so there is no question of the product being subsidised by anyone outside the industry and also no rationing problems. It would be much more difficult, for example, to use a consumers' surplus criterion which would measure the effect of nationalised industries on consumers as a whole; but clearly if it were attempted it would be a different criterion.

There are of course strong arguments against what seems to be the standard opinion that nationalised industries should be run in the consumers' interest; and much to be said for the argument that nationalised industries even in surplus, should maximise their profits, which would mean their purpose is to benefit the public as taxpayer rather than as consumer. Profits would be used either to reduce taxation or to pay for additional government expenditure. The commonest argument against this policy—that the prices of a nationalised industry's goods and services should be kept down as an anti-inflationary measure—is surely almost worthless. Nor is there anything to be said on grounds of allocation for keeping nationalised industries' prices low when private firms' purpose is to keep them as high as is compatible with profit maximisation. (The argument on the grounds that this would mean that the public sector would make exorbitant profits because of its monopoly power we have already examined. (i) Nationalised industries have less monopoly power than is often thought and less than they once had. (ii) Where their profits are judged exorbitant, this is an argument for some restriction on their profit-making ability. It is not an argument against their making profits altogether or against their making lower profits than is usual in private industry. (iii) Since the profits go to the Exchequer there is less cause to be worried about their size.)

That is one reason why it might be judged socially desirable for nationalised industries—and the roads programme—to use a consumers' surplus rate of return. It is a different reason for using it than those which first led to its use in working out rates of return on roads, bridges, etc. If we are to rehabilitate a consumers' surplus rate of return it is worth pointing the contrast between this reason for its adoption and those traditional in roads economics and elsewhere in economics. Dupuit argued for its use on general welfare grounds. He suggested that a consumers' surplus rate of return was better than a private rate of return for all things, since to charge for anything always reduced its utility; unless the charging was perfect discrimination.[1] He argued that ordinarily a charge must be made but that it was possible to adopt this superior criterion for bridges and roads because their variable costs were zero so that there was no problem of rationing the use of a bridge or road.

[1] Dupuit, *op. cit.*

We have seen why we do not have to accept the argument that the criterion is superior on grounds of welfare; why even if we do accept the criterion, we do not have to confine it to cases of zero variable costs, or that if we do insist that price be above a marginal or average cost whichever is the higher, we can use a ' modified ' consumers' surplus criterion. But these things apart, it really is no longer the case that we can regard roads and roadworks as having falling marginal costs. They normally fall over a range but rise again when the weight of traffic increases beyond a certain point. (The concomitant congestion also affects the consumer, and so would also affect a profit-maximising entrepreneur who owned a toll-gate.) Except where they are considerably under-utilised, a reduction in the price of using roads such as is implied by the substitution of a consumers' surplus rate of return for a profit-maximising one does create problems of rationing, the cost of which is ordinarily borne here by the consumer as costs of congestion.[1] Welfare arguments for the superiority of consumers' surplus criterion have been the cause of its becoming impossible to use and also in the end disreputable. Dupuit himself thought that he ought to qualify his consumers' surplus criterion and count in the effect of building a bridge on pre-existing means of crossing the Seine.[2] Since then other protagonists of the consumers' surplus criterion have responded to arguments that they *ought* to count in the effect on producers' surpluses of the commodity itself, substitutes, complements, etc., in ever increasing circles until they have found themselves in principle advocating the impracticable social surplus criterion. And as welfare economists have come to argue that these criteria have no *a priori* superiority, the usefulness of practical surplus criteria other than profit maximisation has tended to be denied or ignored altogether. This is a pity. They have a right to exist. Their use seems to be implied in the adoption of social policies.

But even if the social policy were to be adopted that all nationalised industries and public corporations use profit-maximising rates of return—this would mean a clear distinction between public industries and welfare services—there might still be a reason why the roads programme should use a consumers' surplus rate of return. The ultimate defence for using such a rate of return on roads is that it would be impracticable to use a private rate of return. This is not because such a rate of return could not in practice be estimated but because it would be a political impossibility. Although public opinion is perhaps coming round to the view that a nationalised industry in deficit should rid itself of subsidies and minimise its losses, and may also be beginning to allow the possibility of profit maximisation for nationalised industries in surplus, there is no present likelihood that it will come to think of the building and operating of roads in this way. This would require a revolution in thought and, more important, in practice. It is a simple matter administratively to make nationalised industries maximise their profits, but roads are not organised to make profits, and the changes which would be needed to make this possible are not politically feasible. The argument really is this: any feasible private rate of return which could be devised on road investment would be too small to reflect what is widely thought to be the desirable level of investment in road-building; while a private rate of return which would reflect what was thought to be a desirable level (and allocation) would involve a pricing policy and administrative changes which are politically unacceptable. This is an interesting point because it does suggest that the public is unwilling to accept the weighting which would be given to roads if a feasible private rate of return were used. On the other hand it does seem as if public opinion is more likely to take the straightforward comparison between a private rate of return in private and in nationalised industry as a valid one reflecting what is desirable. (This presumably means that the constraints limiting profits in private and public industry are felt to be socially more acceptable from the point of view of investment governed by private rates of return, than the constraints limiting profit-making on the roads, though these latter constraints are thought to be socially desirable on other grounds.)

---

[1] See D. L. Munby, ' The Roads as Economic Assets ', *supra*.

[2] Dupuit, *op. cit.*, p. 98.

There are three methods by which the roads programme might try to evolve a private rate of return.

(i) *Tolls.* These could be used to this end; but even if we could imagine it politically possible to set up a network of tollgates throughout the country, this would in a small country be an exceedingly costly method of making a profit on the roads. There is perhaps no other commodity of which it can be said that the cost of making a charge is so strong an argument against making a charge at all, by levying tolls.

(ii) *Vehicle and Fuel Taxation.* The present policy is far from using taxation receipts as a measure of the desirability of making an investment. Rather the effect of road improvements and road works is to reduce receipts of taxation by reducing congestion and improving surfaces so reducing the consumption of petrol per mile. It would in principle be extremely difficult to use taxation receipts to decide between alternative road investments if the rate of taxation was not made to vary from the use of one type of road to another and if it were not found possible to alter the system of taxation so that congestion was no longer profitable to the Exchequer.

(iii) *Discriminatory Taxation.* Again in theory it would be possible for a government or highways authority to extract more revenue from road-users, by acting as a discriminatory monopolist than from tolls or vehicle and fuel taxation. Such an authority could act like Professor Pigou's discriminating monopolist of the first degree, making a separate bargain with every road-user, haggling to make him pay the most he would be ready to pay rather than forgo the use of the roads altogether for any purpose: a lump sum, possibly computed annually. This would be a large proportion of anyone's income. It would be quite impossible to practise this discrimination. It would also be absurd to use the power of highway authorities to strike different bargains with different road-users but with less discrimination as a basis for choosing between road investments. Any practical system, if there could be any, would be largely arbitrary in its assessments and would tend to approximate to something like the present system of vehicle taxation.

Therefore it would seem that on grounds of practicability and political possibility, the roads programme cannot use a profit maximising rate of return. Others might argue deliberately that like other nationalised industries it ought not to, for other reasons, if they felt inclined. It is for the first reason that road engineers have felt they have had to adopt another criterion. We can say then that for the nationalised industries to adopt profit maximisation would only accentuate Mr. Glassborow's problem. There is bound to be a break somewhere since, if this argument is right, firms, nationalised industries and the roads programme cannot all use the same rate of return. Private firms cannot be taxed, subsidised or directed so that they use a rate of return comparable to that which the roads programme would use if it used a rate of return at all. However desirable we might think it for private firms to adopt a consumers' surplus rate of return, this is not practical politics. Only the nationalised industries would seem to have a choice. From one point of view it might well be that the Spirit of the Times would like nationalised industries to behave like private industries—and I would not myself dissent from this view—by and large, maximising their profits. This is what comes, or is starting to come, into its mind's eye when it thinks about the relation between private and public industry, and private and public investment. It sees no reason for drawing a distinction. But when it turns its head and looks in the other direction and compares one nationalised industry, British Railways, with investment in roads, it does not know what to think. From this point of view there is something clearly wrong in British Railways using a private rate of return—and it would still seem wrong if British Railways purged itself of its present haphazard accumulation of elements of subsidy—and the roads programme using its approximation to a consumers' surplus rate of return. (We cannot conclude necessarily that this exaggerates the returns on road investment because of the way in which consumers' surplus is estimated there.)

There is a way out of this dilemma in principle which has been indicated already. This is that the government should use the rate of return on nationalised industries it thinks fit, and if this should be different from the rate of return used on roads, then it should make a decision of principle how these different rates of return should be weighted in allocating investment funds between the roads and nationalised industry. However, this is a type of decision which a government cannot at present be expected to make. It lies beyond its experience. The only practical way out of this dilemma is for society to decide which comparison matters more: the comparison between British Railways in particular and investment in roads; or the comparison between British Railways as a nationalised industry and the rest of industry, private and nationalised. Surely there can be no doubt that the first comparison matters more. It matters more that the most efficient allocation of resources should be made between road and rail, which would be more easily achieved if the same rate of return should be used for both, than that in the name of being businesslike British Railways should be expected to show an ordinary rate of return on the use of its capital. It follows that unless there are good reasons for treating all nationalised industries in the same way, which I do not myself believe there are, there is a case for calculating a different, consumers' surplus rate of return on railway investment from that used in other nationalised industries. (The important comparison for instance for the National Coal Board, the Central Electricity Authority and the Gas Boards is surely with oil and this suggests that the best compromise solution would be a private rate of return.) I do not see why this innovation should worry us. There will, of course, be a problem in deciding how funds should be allocated between consumers' surplus public enterprise and firms and industries using a profit maximisation criterion, but the break there is surely less serious than one between the road and rail criteria; and can be less hazardously bridged by rough value judgments. Although it is possible to suggest modifications in detail of the way in which the Road Research Laboratory calculates its rate of return, it is this we must take as the fixed point and redefine the rail rate of return so that it is comparable with it. This means we would use the same approximation to the consumers' surplus rate of return on rail investment; and that we should charge in both cases a price to cover long-run marginal costs. There are of course difficulties which arise because the railways provide vehicles as well as track. These are difficulties of detail. But comparability in the sense here used, given the difficulties of the situation, seems a sensible though unfamiliar thing to do.

There is a last point. In estimating the rate of return on road investment, the Road Research Laboratory takes into account the benefits and costs which accrue to other roads from the building of a new road, but does not consider the effect of a new road on the railways (or on canals and airlines for that matter). This one could broadly describe as a competitive policy; the justification being that though both are government-owned, roads and rail must be in competition to achieve the best service at the lowest cost. This is clearly desirable if it is not possible to provide as good a service as cheaply by administration and accounting methods rather than by competition. The alternative policy is what is sometimes loosely called co-ordination. We can give this a stricter sense by saying that it means in the context of investment that whenever the rate of return is worked out on an investment in a road, the effect of the investment on the relevant railway services whether plus or minus is allowed for; and conversely with an investment in rail. At the present moment, when our railways are under-utilised and our roads insufficient, this would seem a more sensible interim policy than competition, whatever we may decide when our roads become more adequate—that is, when the rate of return on new road projects declines towards the socially desirable rate of profit, however calculated, in industry. If a co-ordinated investment policy were practised initially this would mean that roads would be built first where alternative railway services were least adequate.

Jesus College,
Oxford

INVESTMENT IN ROADS IN UNDERDEVELOPED COUNTRIES[1]

By E. K. HAWKINS

The one sure generalization that can be made about the underdeveloped countries is that investment in transport and communications is a vital factor. Economists have accepted this since the time of Adam Smith and it rarely seemed necessary to back this judgement with calculations of the rates of return obtainable from investments in roads. Unlike railways and canals their profitability could not be assessed directly. In any case, it was felt that they were so much a part of the essential structure of public works required in a country as not to need more formal justification.

When the first development plans were drawn up for British dependent territories during and after the second world war roads occupied a prominent place in them and have continued to do so. The reasons for this are mixed: it was widely felt that they had been badly neglected in the pre-war years of depression and retrenchment.[2] For the first time money and resources were available for the completion of road development plans made many years before. The need for them was even greater, now that the full implications of the value of road transport had come to be appreciated by all groups in society. There was an element of relief amongst those who had to draw up such development plans that an item of expenditure was so readily to hand which was not likely to be controversial. There may also be something in the argument that roads were included, to the extent that they were, precisely because there was no easy way of demonstrating at a later date that the expenditures had been wasted because the projects chosen had made a ' loss '. Such fears certainly ruled out some agricultural and industrial projects and a number of possible railway extensions.

As long as funds for development were plentiful there were few critics of expenditures on road systems and substantial mileages were built and reconstructed in many British and former British territories. As funds became scarcer and there were more claimants for them it became necessary to defend the case for spending more on roads, especially as the more obvious projects on which all could be agreed had been completed. Both engineers and administrators became interested in ways of assessing the returns to be obtained from such expenditures. In countries where there were rival forms of transport with claims on government funds the question was more acute and it was more important than ever that the funds should be allocated in a rational fashion between the competing projects.

---

[1] I am indebted to Mr. Paul Streeten for his comments on an earlier draft. The final preparation of this paper was assisted by the University of Sheffield Research Fund.

[2] In some cases this neglect had been deliberate, in order to protect railway systems from competition.

These questions have been widely considered elsewhere in connection with road development in more advanced countries. It might be thought, therefore, that it is simply a matter of adapting the techniques used to the special circumstances of underdeveloped countries. This is so, in fact, but the process of adaptation involves a marked change of emphasis and is worth discussing in some detail.

The benefits that flow from an investment in road construction or improvements take three forms. First, there are those direct benefits that go to the users of the road. Secondly, there will be benefits that are distributed more widely than to the direct user of the facility. It is usually these wider effects that economists have in mind when talking of the stimulating effects of road investments on economic development. They can come only in two ways: by the diffusion of the direct benefits received by the user of the road, or through external economies made possible by the road, but not attributable to particular users. Thirdly, there may be gains to the public authorities in the form of lower costs of maintenance for the whole road system and, possibly, also additional tax revenue from the greater use made of the roads by users. Care must be taken not to double-count this last item, since it derives from the benefits already included under the first heading above.

In the second case the exact way in which the benefits get diffused will depend upon the reaction of those users who receive them. The most common form that this could take would be for the reduction in the costs of operating road vehicles to be passed on in the form of lower rates and charges. The manner in which this would happen and the speed of the process would depend upon a number of factors which are often ignored or taken for granted. It makes considerable difference to the result, for example, if the road transport is provided by owner-operators, or by entrepreneurs operating transport on their own account in the course of their business. In the former case the benefits of the new investment reach the user directly in the form of lower costs and larger profits. If road transport services are provided by specialists offering their vehicles for hire the immediate benefits will accrue to them. The extent to which they will then be passed on in lower rates for the customers will depend upon the way in which the industry is organised. Where competition between transport firms is effective it is very probable that the lower costs of operation will be passed on in the form of lower rates.

In cases where road transport is provided by firms who operate under monopolistic conditions there may not be the same incentives to pass on lower costs in the form of lower rates. The benefits of the investment will then tend to ' stick ' with the transport firms and boost their profits. They may still ultimately be diffused through the economy when these profits are disbursed, but the manner in which this will happen will clearly differ from the first case where there was an immediate and less concentrated distribution amongst a large number of road users. In some underdeveloped countries road transport firms are owned by foreigners. Where they enjoy monopolistic conditions the benefits of road improvements may flow overseas in the form of higher profits remitted to foreign shareholders.[1]

When one turns to the wider benefits attributed to investments in roads in underdeveloped countries there are difficulties in defining exactly what is meant, over and above the diffusion of the direct benefits just discussed. There is no doubt, however, that it is these wider, external economies that are thought to be the most important results of such investments.[2] Such an assumption requires rather more justification than it has received in the past. One possible kind of result to be expected of road improvements will be that new users will travel who did not do so before. This ' generated traffic ' is an indirect effect only in the formal sense that more travel is undertaken now that transport costs are lower. Unless the road is a new one it is difficult to measure the extent of this induced effect. Traffic is usually growing, even on unimproved roads, so that it is difficult to assess which users are now travelling who did not do so before.

Another indirect effect of road improvements will be the net decrease in the costs of accidents to the community. In more developed countries it seems certain that better roads decrease the incidence of accidents per vehicle-mile. It may also be so in underdeveloped countries, although the data are too sketchy, in many cases, to allow firm deductions to be drawn from them. There are some considerations, however, which would indicate that road improvements might increase the burden of accidents in underdeveloped countries.

The unimproved roads in such countries are usually so indifferent in quality as to instil caution into the boldest drivers. Average speeds tend to be low, despite the light traffic that is the situation on most roads. When the roads are improved, the low traffic density permits vehicles to travel at relatively high speeds. The result is that accidents, when they occur, are often more serious in loss of life and damage to property. They may also be more frequent per vehicle-mile travelled, although the statistical evidence is sketchy on this point. In any case this possibility is only likely to be true for comparisons between improved and unimproved roads, given the same traffic flow on each. When traffic is increasing, the rate of increase of accidents will almost certainly be lower on the improved roads.

The burden of injury and loss of life from road accidents is very heavy in underdeveloped countries. The skilled and educated part of the population is a small proportion of the total; it is these people who have more chances of being concerned in accidents, since they travel more than the average and they are the vehicle-owning section of the community. If they are killed or disabled, the loss to the community is a heavy one.[1]

---

[1] The most likely cause of monopolistic conditions in road transport would be restrictions on entry into the industry. These can be the result of official licensing procedures, although the same official intervention can be used to control rates, services and the level of profits.

[2] For example: ' Overseas the benefits from new road construction are almost entirely in the form of new development, from the traffic which the new road will generate.' R. S. Millard, ' Road Development in the Overseas Territories,' *Journal of the Royal Society of Arts*, 1959, page 275.

So far two indirect benefits of road improvements have been mentioned, possible reductions in the costs and frequency of accidents and the growth of induced, or ' generated ' traffic. The first of these can be described as an external economy, in the sense that the benefits are ' dependent on the general development of the industry ',[2] and it would not be possible for the owners of the road (in this case, the state), to appropriate them to its own revenues by means of some system of taxes or tolls. In addition, the benefits do not accrue solely to one individual, or group of individuals. This is not the case with induced traffic, however, which is the result of an increase in the demand for the services of the road following a reduction in the costs of those services; the benefits go, in the first instance, to individual users of the road.

The most generally recognised way in which transport improvements stimulate development is by creating new or better means of communication; it is clear that ' generated ' traffic is a sign that this has happened. The mechanism, however, operates through a reduction in transport costs and a consequent widening of the market. The following example, taken from the *Report of the East African Royal Commission* 1953–55 illustrates the process very well:

> ' In a recent study of a coastal tribe it has been pointed out that the absence of a road to bring crops from the fields into the coastal villages necessitates the carriage of grain by head load or donkey. The cost of transport by head porterage was from Shs. 8/- to Shs. 12/- a ton-mile, and by donkey Shs. 7/- to Shs. 12/-. By lorry it would have been Shs. 3/50 to Shs. 5/-. Transport costs at Shs. 10/- a ton-mile would, if the cultivator's plot is ten miles from the village, form about 20 per cent of the maximum gross return, and if the plot is 15 miles from the village this figure would rise to 30 per cent. Maize could, therefore, be farmed only within a few miles of each coastal village. . . . When in 1953 a road was completed, it was found that land had been cleared and planted as far as 18 miles away in anticipation of the completion of the road.' (p. 120)

In this case the developmental effects operated through the anticipation of a new means of transport being available which would halve the costs of carriage per mile. Those who had previously used draught animals would either have to obtain lorries to gain these benefits, or they would have to hire the services of firms of transporters. In this second case, much the more likely in the above example, it is important, as already pointed out above, that the low costs of vehicle operation be passed on to the farmer in the rates charged.

The community may place some value on the mere existence of a new or improved road, even if there were no cost-reducing effects of the above type.

---

[1] The same argument can be used to show the high potential value of the time saved by travellers as a result of road improvements. It was so used in the nineteenth century: ' . . . and as travellers in general belong to the superior and the more intelligent classes their time is proportionally more valuable.' D. Lardner, *Railway Economy* (London), 1850, page 13.

[2] A. Marshall, *Principles of Economics*, (London), 1920, page 266.

Where roads and communications are bad they are often uncertain and erratic; when emergencies arise and patients have to be rushed to hospital, or police and administrators have to travel quickly to an area, certainty of communications is an important consideration. In more developed countries it is taken for granted: in underdeveloped countries, especially those in the tropics, the sudden failure of a gravel or earth road, usually in wet weather, imposes considerable costs on the economy, although they cannot always be measured.

This last case can also be described as a true external economy, although it is true that, in a formal sense, the immediate benefits are enjoyed by those who make use of the better roads in emergencies. Where they are agents for the whole community, however, such as administrators or the police, the benefits also accrue to the community as a whole. In addition, the certainty of communications reduces costs throughout the whole economy, regardless of the benefits to those who actually make use of them. It is not necessary, for example, to make such widespread arrangements for the maintenance of law and order, or to hold large stocks of fuel, foodstuffs and medical supplies locally, if one can be certain of obtaining aid, if needed, from other parts of the country.

It has thus been possible to identify two external economies arising out of road improvements. One is the lower social costs of road accidents and the other is the value that can be placed on greater certainty of communications. All the other wider effects of investments in road improvements are the results of the direct benefits being diffused through the economy. If the total value of road investments is to be correctly assessed, this could only be done by measuring the ultimate annual increase in the national income that can be attributed to the investment. In other words, one would need to know the value of the ' super-multiplier ', so that both the increases in income associated directly with the investment are taken into account and any further rounds of induced investment as well. For practical purposes this is usually impossible, even for developed economies with sophisticated statistical services. In underdeveloped economies not only is the evidence not usually available, but it is impossible in principle to sort out the effects, *ex post*, of investments in roads and transport from those of other investments. There is everything to be said, therefore, for limiting the assessment of the possible returns from road investments to those direct benefits that can be measured. At a second stage, it may be possible to take into account some of the wider effects upon incomes and trade, as well as any non-economic factors.

The estimation of the direct benefits of an investment in roads can be done by means of the cost/benefit method which has been widely used elsewhere. The adaptation to the conditions of underdeveloped territories requires a change of emphasis. The method consists of comparing the prospective yields from an increment of capital expenditure upon roads with the net increase in the costs to which that expenditure gives rise. Alternatively, the sum of all future prospective yields, net of additional expected recurrent costs, can be compared with the present capital value of the scheme. This will give a rate of return over cost for different schemes, allowing them to be ranked in order of expected profitability.

For some purposes only the first and simplest method need be used; that is to say, the expected immediate annual yield from the investment is compared with the additional annual costs incurred. This can be done without difficulty from existing statistics of traffic and costs. If the calculation shows a positive return the investment has passed what might be called the ' threshold ' test. This is a useful test to apply in order to produce a ' short list ' of possible investment projects when funds are limited: there has to be a very strong case in underdeveloped countries before it is worthwhile starting a project which will give a negative yield in the first year.

The ' threshold ' test is not a sufficient criterion, however, because roads have long lives and it is the yield over the whole of the life, suitably discounted, that matters. The rate of return over cost will take account of all future expected yields and costs. It can be used to distinguish between projects whose ' threshold ' yields may be the same, but where the future prospects may differ considerably. It will also deal with any case in which the rate of return over cost is positive, but the ' threshold ' yield in the first year is negative. It may be necessary to construct a road which has a much greater capacity than can be justified by the existing traffic; the immediate yields may be negative, but the growth prospects will make them positive at a later date.

A more important use of the rate of return over cost arises when the problem is that of deciding what quality of road to build or reconstruct, given certain expected yields. This type of decision often has to be made in underdeveloped countries, but it has received much less attention from economists in advanced countries. Perhaps the reason is that it tends to be regarded as a technical problem to be solved when the specifications are drawn up by the engineers. The range of choice of types of road (and hence the range of costs) is often relatively small, so that very marginal changes in investment programmes are involved.[1] For poorer countries, however, the variation in the cost per mile between high-cost and low-cost roads may result in a substantial change in the public works budget. This only reflects the familiar generalization that the cost of public works represents a proportionately heavier burden on underdeveloped economies than it does for more advanced countries.

The calculation of the rate of return over cost involves estimating the present value of all direct benefits expected from the investment and comparing them with the present capital cost of the scheme. The latter presents no problem, so the main concern is the correct estimation of future expected benefits, the choice of an appropriate time period over which to measure them and the rate of interest to be used as a discount factor.

The benefits to be evaluated consist of (*a*) reductions in the operating costs of vehicles per mile, (*b*) the value of time saved by vehicle operators and transport users; (*c*) the net effect on the frequency and cost to the community of road accidents. They must be assessed with suitable assumptions about the rate of growth of traffic, both that which would have occurred in any case and the additional traffic ' generated ' by the investment.

---

[1] The high opportunity cost of land in developed economies, which is invariant between different types of road, has much to do with this situation.

The work done in the United Kingdom stresses items (*b*) and (*c*)[1], whilst (*a*) is regarded as being of subsidiary importance. The emphasis placed on (*b*), the value of time saved, derives from the use made of traffic engineering studies which relate the traffic flow along a particular road to the average speed of that traffic and the components of the flow. The virtue of this approach is that it uses relatively simple arguments based on empirical relationships to obtain measures of the way in which the cost of transport rises as the limits of the capacity of a road are approached.

The basic argument can be summarized as follows: the driver of a vehicle will choose to travel at the speed best suited to the conditions of the road, the type and condition of the vehicle and his personal preferences. (These may include an explicit recognition of the fact that the cost per vehicle-mile rises as the average speed increases). In conditions of low traffic density the driver's chosen speed will not be affected by the presence of other vehicles on the road, since there will be ample opportunities for passing or being passed. In these circumstances the pattern of vehicle speeds will represent what is known as the ' free speed distribution '.[2] As the traffic flow increases there will be a fall in the average speed that drivers can maintain. It has been found that there is no relationship between the average speed of traffic and the size of the traffic flow until the flow reaches a certain minimum critical size.[3] Beyond that point the average speed of the traffic is inversely related to the size of the traffic flow. In economic terms, in the range where speed is unaffected by the amount of traffic the cost of transport per mile for the individual operator depends solely upon his own choice of conditions. Where the flows are larger, these costs come to depend also upon the number of vehicles using the road. In other words, the social cost per mile begins to exceed the private cost. The addition of one more vehicle to the traffic flow increases the total cost of transport along that route by more than the costs met by the new operator, because his vehicle increases the costs for everyone else on the road.

The nature of this increase in costs is the result of two opposing effects. The enforced fall in the speed of vehicles will reduce those costs (such as fuel consumption) which vary with speed. On the other hand drivers are now losing time by travelling slower than they would freely choose to do. The net effects of these two tendencies will depend upon the value placed upon the time lost through increasing congestion. It has been shown that, with very large traffic flows, the aggregate value of the time lost can be considerable.[1]

---

[1] See W. H. Glanville and R. J. Smeed, ' The basic requirements for the roads of Great Britain ', *Proceedings of the Conference on the Highway Needs of Great Britain*, 1957; T. M. Coburn, M. E. Beesley and D. J. Reynolds, *The London-Birmingham Motorway, Traffic and Economics* (London), 1960.

[2] M. Beckmann, C. B. McGuire and C. B. Winsten, *Studies in the Economics of Transportation* (New Haven), 1956, page 22.

[3] J. G. Wardrop and J. T. Duff, ' Factors affecting road capacity ', *Proceedings of the International Study Week in Traffic Engineering* (London), 1956.

For countries such as the United Kingdom, where traffic congestion is a serious problem, the above approach suggests that the role of investments in road improvements should be to alter the relationships between speed and flow so that the improved road will either carry the same flows as before at higher average speeds, or larger flows at the same average speed. The traffic characteristics of most underdeveloped countries do not correspond to this kind of situation. Apart from the congestion experienced in the large cities, such as Lagos, Nairobi and Salisbury, traffic flows tend to be small.[2] It seems that the yield from road investments cannot be discussed mainly in terms of the costs of congestion and the value of time saved. Instead it is the reduction in the costs of vehicle operation that is the most important source of benefits to be obtained from expenditures on new and better roads in underdeveloped countries.

The most important source of these reductions in costs comes from improvements to the surfaces of the roads over which the vehicles run. Road improvements in the United Kingdom normally provide roads which have better alignments and much larger capacities; the quality of the riding surface is not likely to be markedly different from that of the older roads. In underdeveloped countries, however, the object of investments will often include the replacement of an earth or gravel surface with a permanent hard surface, such as bitumen.

Very fine roads can be constructed with gravel surfaces[3] whose riding qualities are in no way inferior to hard-surfaced roads. They cannot be kept in the best condition for long, however, without extravagant attention and they deteriorate at an alarming rate, even under relatively light traffic. In dry weather they become dusty and corrugated[4] and in wet weather they may become impassable. The capacity of a gravel-surfaced road is much below that of a hard-surfaced, ' all weather ' road of the same dimensions. Speeds are lower, wear and tear on the vehicles is greater and the costs of maintenance, whilst lower than those of a bitumen road, increase at a rapid rate once the traffic flow reaches moderate levels.[5]

Vehicle operators much prefer hard-surfaced roads because of the big difference in the life of vehicles, compared with gravel roads. This fact is widely recognised by all concerned with vehicle operation in underdeveloped countries. It is very difficult to demonstrate the differences in economic life statistically, although many operators of commercial vehicles have some idea of the differences. What is required, strictly speaking, is information covering a lengthy period for two sets of identical vehicles, one set operating on unimproved gravel roads and the other on reconstructed roads.

---

[1] Glanville and Smeed, *op. cit.*

[2] As far as the writer is aware no speed/flow relationships have been published for underdeveloped countries. Experiments made by the writer in Uganda in 1957–58 indicated that there was no relationship between speed and flow on roads outside the few urban areas.

[3] This is particularly true of the lateritic gravels which are widespread in Africa. They are known as Laterite in West Africa and as ' Murram ' in East Africa.

[4] ' . . . With the passage of about 100 vehicles, a newly levelled gravel road can develop a surface like a huge washboard with regular transverse waves, 1½ in. or more deep, some 2 to 3 ft. from crest to crest.' R. S. Millard, *op. cit.*, page 277.

[5] A rough ' rule of thumb ' has been used by engineers in African countries to the effect that once traffic reaches 300 vehicles per day, it is less costly, in the long run, to provide a hard surface, because of the difficulty of maintaining gravel roads.

Results from statistical enquiries of this kind are not yet available. However some estimates can be made which show the order of magnitude of the reduction in operating costs. The writer has collected information in East Africa which shows an average difference of 3d. per vehicle-mile between gravel surfaced roads and those with ‘ all weather ’ hard surfaces.[1] Given a traffic flow of 300 vehicles per day, this would amount to approximately £1,400 per year. The greater part of this reduction in operating costs comes from the lower depreciation charge per mile, as a result of the longer economic life of the vehicle. The relative importance of the different cost items is shown in the following Table:

Percentage distribution of total average operating costs per vehicle-mile for selected vehicles in Uganda, 1957

Type of Vehicle	*Carrying Capacity (tons)*	*Fuel and Oil*	*Tyres*	*Repairs and Maintenance*	*Wages*	*Taxes and Insurance*	*Depreciation*	*Total*
Pick-up	0.75	28	5	24	18	5	20	100
Lorry	1.5	27	4	14	24	6	25	100
Lorry	3.0	37	2	14	21	8	18	100
Lorry	5.0	31	7	14	10	7	31	100

Source: Data collected by the writer.

While the main benefit of road improvements will take the form of lower costs of vehicle operation per mile, there will also be some savings in time to take into account. For example, the reconstruction (including re-alignment) of one main road in Uganda reduced the average journey time from five hours to two. It has been argued above that these savings will not be as important as they are in more developed economies, but the exact extent of them will depend upon the condition of the road before and after improvement and the level of traffic.

Fundamentally, these savings in time achieve the same economic purpose as the reduction in operating costs. Both result in economies in capital used in road transport, by enabling the same investment in vehicles to do more work. The savings in time permit more vehicle-miles per day or per year; the reduction in operating costs is due to lower depreciation per mile, since vehicles have longer lives on the better roads.[2]

It is possible, in principle, although often difficult in practice, to place appropriate valuations on the reductions in vehicle operating costs, the savings in time and the net effect upon the costs of road accidents. Once the rate of growth of traffic has been assessed, it is then possible to estimate the present value of all the future expected benefits from an investment in a road project. Before this can be done a decision must be taken as to the period of time over which these benefits are to be measured; in other words, what is the expected economic life of the investment? Secondly, what rate of interest is to be used as a discount factor? This is also relevant to the calculation of the benefits side of the equation. In order to get the net annual expected benefits it is necessary to deduct the additional annual costs which result from the investment. These consist of the net additional maintenance costs per mile of road and, secondly, of an interest charge on the capital cost of the project. The rate of interest chosen for this should clearly be the same as that used as a discount factor. In theory it should be the highest current rate of return obtainable on capital elsewhere in the economy.

---

[1] It is interesting to note that Millard puts the possible saving per mile at twice this amount, *op. cit.*, page 275.

[2] See Coburn, Beesley and Reynolds, *op. cit.*, page 47.

The choice of a time period over which to calculate the present value of the investment presents more difficulty. One solution is to treat roads as perpetual assets; the discounting process will reduce the more distant yields to negligible amounts, in any case. Where capital is scarce, then the discount factor will be large, having a very big effect upon the distant expected benefits.

While it is true that there are some parts of a road that will last for ever, provided that proper maintenance is carried out, there are many cases in which roads need to be rebuilt, abandoned or replaced in entirety. This is often hidden in practice by the problem of distinguishing between routine maintenance and reconstruction and improvements. This is a difficulty which occurs in all road programmes because, for engineering purposes, the two are often carried out together and can be viewed as substitutes. What happens is that the roads ' wear out ', or come to the end of their economic life and this is signalled by the real costs of maintenance per mile rising over time. At some point it becomes more economic to reconstruct and improve them, rather than try to maintain the road in its original form. This situation is nearly always connected with a rising trend of traffic as the road reaches the limits of its capacity.

There may also be cases, however, where a road becomes obsolete because of a change in economic circumstances in its immediate hinterland, or because of economic development elsewhere. There is also an associated but essentially different type of obsolescence which arises from the possibilities of technical advances in the art of road construction and maintenance. Cheaper and more efficient ways of construction and maintenance may be developed which can only be applied by reconstructing existing roads, thus terminating their life.[1] Both types of obsolence are likely to be more important in underdeveloped countries than elsewhere: the lack of knowledge both of future regional and local patterns of development and also of possible technical advances will ensure this.

---

[1] A good example of a technical advance of this type is the development of ' soil stabilisation ' techniques. These permit the construction of roads with good load-bearing properties without the use of an expensive stone base. They can only be adopted, however, by tearing up the existing road surface and rebuilding it.

All these considerations make it desirable that any calculations of yields from road investments should be related to a definite time period for the life of the asset. Even if the economic life of a road is not known—and there are virtually no data available on this subject—it is desirable to assume one for planning purposes. Where capital is scarce, ' it is not sound or humane policy to burden the present overmuch for the sake of the future . . .'[1] This can happen all too easily by building roads to standards that are too high in terms of the expected yields from them. It is frequently necessary to choose between road projects of different standards and costs. Such choices make economic sense only if it is known, or can be assumed, that the more expensive project will have a longer life; then the choice between projects will depend on the expected yields, the rate of interest and the costs of construction. All these can be summarized in the rate of return over cost, which will show which project is the more profitable.

This example is mentioned because it is one that arises very often in road development programmes. Given a limited budget, is it better to develop a big mileage of low quality roads, or to concentrate on providing a smaller mileage of high quality roads? If the necessary facts are available it is clear that an unequivocal answer can be given to this problem in any particular set of circumstances by the use of the type of calculation outlined above.

A variant of the above example concerns the question as to whether it is worthwhile providing a gravel road with an all-weather hard surface if it is not to be reconstructed and realigned first. Engineers often argue that to do so is rather like putting new wine into old bottles. Much will depend upon local circumstances, because the costs of road construction vary with the locality. It has been shown above that the provision of a hard surface is likely to lead to considerable economies in the capital invested in vehicles. Realignment and reconstruction, on the other hand, lead to gains *via* higher vehicle speeds, fewer accidents and a bigger capacity for the road. This is also a case where the right economic choices would be facilitated by bringing all the relevant information under review in the calculation of the rate of return over costs.

It is probably superfluous to add that the object of exercises of this kind is not to forecast the future with accuracy. The aim is to check whether rival investment proposals are likely to show an economic return when all the present knowledge and best possible assumptions about unknown factors have been taken into account. Too often government planning in underdeveloped countries is based on assertions of a general nature. The logic of the method outlined above forces all the facts and assumptions out into the open, where it can be seen whether they are consistent or not.

University of Sheffield.

---

[1] J. M. Keynes, *Indian Currency and Finance* (London), 1913, page 170.

ROAD PLANNING: SOME LESSONS LEARNED IN BRITISH NORTH BORNEO

By J. R. SARGENT

Some might think Blackwood's magazine a more appropriate journal than the BULLETIN for reminiscences of travel in Borneo. But the intrepid economist hacking his way through the jungle hears many sounds which resolve themselves to his surprise into economic arguments. I have therefore picked out some of these and set them down, in the hope that they may be of interest to subsequent explorers in similar territories.

In a country like Borneo the most significant returns to capital spent in building roads are likely to be found in their effect in increasing the area of cultivable land. From a somewhat crude comparison, district by district, of the mileage of new road constructed and the expansion of the cultivated area over the previous six years, my own guess was that an average of 250 acres a mile had been gained. If this were planted with rubber, which shares with timber the status of the Colony's most important product, the increase in the national income would represent an annual return of more than four times the capital cost. A gain of this size could hardly fail to dwarf any conceivable benefit from savings in vehicle running costs—attributable to greater speeds or reduced maintenance on improved surfaces—or from economies of scale due to wider markets—in a country of 400,000 people, where industrial enterprise is up against competition from Hong Kong with its plentiful skills and swelling reserves of low-paid labour—or from replacement of existing means of transport such as coastwise shipping or (in part of the Colony) rail. Nevertheless a claim of an annual return of 400 per cent is bound to arouse suspicion. In this case it is built upon things which are in the very literal sense of the word variables, such as the price of rubber (over time) and the yield per acre (according to whether it is cultivated by Chinese smallholders, large estates, or the indigenous population); although middling values have been taken for these. The estimate also assumes that the road required to open up the land is of earth only and is unmetalled, although it satisfies engineering standards, with regard to alignment and drainage, laid down for a road of that type by the Colony's Public Works Department. The view is sometimes held that the success of a road in opening up land is a function of the standard to which it is built. I was able to find no clear support for this view in the results taken district by district, and consequently have seen no reason to base the estimated return on a capital cost higher than the minimum.

To my mind, however, there is less reason to be suspicious of the ingredients of the estimate than cautious about its interpretation. In the first place it assumes that the mile of road which opens up the 250 acres is being pushed outwards from the existing network.[1] If it is successful, it will involve a greater pressure of traffic on the latter, the cost of which should be properly taken account of. But suppose that some land is found of a higher fertility than any on the existing margin, which can only be reached by extending one of the existing roads to reach the fertile area through country which is economically useless. A given capital expenditure on roads will then either bring fewer acres into production at a higher net output per acre, or more at a lower. This choice recently arose in Borneo owing to the work of government soil scientists, who discovered a substantial area of fertile soil in comparatively inaccessible parts of the Colony. The solution to it is not simply ' Calculate! Calculate! '; and this brings me to the second reason for caution.

This is that it is not roads that open up land but men. To estimate the marginal return to capital spent on roads by multiplying the estimated marginal acreage by an estimate of net output per acre implicitly assumes that the necessary labour can be conjured up without a decline in output elsewhere. In most of the under-developed countries of South-East Asia it would be reasonable to hope that this would be so. But conditions in British North Borneo are an exception to those which generally prevail in South-East Asia; it is not an over-populated country. Consequently larger additions to the cultivated acreage promised by the one alternative might not be able to be realised, except at the cost of a bigger reduction of output elsewhere than the other alternative would involve.[2] Could not any shortage of labour be overcome by immigration? There is certainly no lack of willing immigrants, particularly since discrimination against the Chinese has erupted in Indonesia and elsewhere. Here, however, the economist runs his head into the political wall of suspicion of Chinese immigrants.

That the supply of labour has in fact been important in British North Borneo was suggested to me by another crude statistical comparison. This set the increase in the cultivated acreage per mile of new road constructed since 1952 in each district—that is, a measure of the relative success of its new roads in opening up land—against an estimate of population growth in the district.[3] The districts with the highest population growth were found to be those in which the success of new roads in extending the cultivated area had been most marked, and the least successful to be those which had

---

[1] That is, the figure was arrived at from a study of a past period when road development had mainly taken this form.

[2] This is an example of ' factor myopia ', or disregard of the complementarity of the factors of production, of which another is to be found in the attitude of agricultural officials to timber production. They argue that the latter should not be encouraged owing to the low productivity of land which characterises it, this being due to the fact that an acre logged today cannot be logged again for (in Borneo) 80–100 years. From the standpoint of the productivity of labour, however, the relative advantage of agriculture is not clear.

[3] Only the natural rate of increase (births minus deaths) was known for each district. Internal migration was not. An indication of the direction of the latter, however, was derived by applying the natural rate of increase for the Colony as a whole since 1951 to the 1951 Census figures for each district. This yielded an ' expected ' natural increase for the district. If the actual increase there was found to be greater than the ' expected ' increase, this was interpreted as due to a net influx into the district, and vice-versa. The results for the Colony as a whole yielded a fairly consistent picture of population movement, mainly from north to south, and from remote districts towards the major centres.

had the lowest population growth. This provided one working general criterion for the allocation of road-building funds. It also helped to resolve the problem of choice to which I have referred, in favour of developing access to the fertile but remote area recently discovered, rather than generally extending the margin of cultivation on less fertile land lying close beyond the limits of existing roads. For the larger acreage which the latter alternative would theoretically have opened up would not in practice have been developed owing to the larger increment of man-power which it would have required.[1] In addition, the more fertile land of the former alternative was fairly close to one of the districts of the Colony which, in exception to its general condition, happened to be over-populated. It also lay athwart what appeared to be the main stream of population movement within the Colony.[2]

The existence of over-populated areas within an under-populated country raises a general point concerning what might be called 'Nurkse surpluses'. The disguised rural under-employment to which he drew attention[3] may not be a general characteristic of an under-developed country, but if transport facilities are limited it can appear on a substantial scale in certain areas. Investment in roads (or other forms of transport) then finds its social return in its effect in equalising the marginal productivity of labour in different areas. But disguised under-employment also appears in a seasonal form where the local population is largely devoted to cultivating a crop such as rice. Here the social cost of road construction can be kept down if local labour can be mobilised for the purpose in the off-season. Yet this can involve an economic choice. For the skill, tact and knowledge of local conditions and attitudes, upon which depends the success of such a mobilisation of seasonally under-employed labour, is most likely to reside in the district administration. On the other hand, the efficiency with which the work is planned and executed is likely to be greatest if it is left to the professional engineers of the Public Works Department, whose knowledge of local conditions is usually less comprehensive. In principle, co-operation between the administration and the professionals should enable both to make their peculiar contribution; in practice it is more likely to be the one or the other. Moreover, if the main effort of road construction is to be confined to the off-season, there is a risk either of capital equipment standing idle in the period from ploughing to harvest, or of it having to be moved to other purposes; and both of these disadvantages are costly.

Disguised under-employment, however, is not the only form which disguised savings can take. British North Borneo shares what appears to be a common feature of Eastern and African countries, namely a number of shops and other retail outlets which seems excessively large in relation to the size of the product available for distribution, and involves a correspondingly large ratio of stocks to sales. This phenomenon might be welcomed by

---

[1] Although this would have happened in the long run owing to the natural growth of the population.

[2] See footnote 3, p. 372.

[3] *Problems of Capital Formation in Underdeveloped Countries*, Chap. II.

Messrs. Bauer and Yamey, whose ' . . . criterion for judging economic development is the widening of the range of alternatives open to people as consumers and as producers '[1] But it is difficult to believe that the size of stocks represents either a social optimum or a private one. If stocks were smaller, private profit would be greater; but this possibility is not exploited because loans and other monetary embodiments of wealth are either not available or not liked. This is a general problem.[2] Its immediate relevance is that one of the forms that hoarding takes is the possession of motor vehicles. I carried out a census of traffic on one of the most heavily used stretches of road in Borneo, from which it emerged that the average utilisation of capacity was below 50 per cent. It is possible that this low figure could be explained by an unbalanced distribution of traffic over time and space; for the railway, which has almost a complete monopoly of transport in another part of the Colony, works to a load factor which is not much better. But I suspect that hoarding would be a substantial ingredient of the explanation of the excessive ratio of vehicles to passenger- and ton-miles although the desire for status-symbols may be entangled with the sheer wealth-holding motive. The inflated number of privately-owned vehicles creates a problem for the public sector, carrying for hire and reward under a system of regulation and public obligation imported from the United Kingdom. Some amendment of this system will clearly be necessary, if public transport is to survive, although its demise would perhaps help to ease the general excess of vehicles. On the assumption that it should survive, it seemed best to recommend the adoption of a differential licensing system, under which private vehicles would be licensed for use at a fee from which operators accepting public obligations would be exempt.

Road planning is a subject on which the views of economists will often collide with those of engineers. Nowhere is this more likely than on the question of low-cost roads. British North Borneo is not alone in possessing a number of such roads, constructed out of earth and with the minimum of equipment under the enthusiastic but amateur direction of district administrations. In equatorial rains, however, these roads are liable to disintegrate unless, at the time of construction, careful attention has been paid to their alignment and drainage. This leads to pressure from the professional engineers to the effect that earth roads should not be built except to what they regard as the proper standard, and this is naturally a more costly standard. They justify the extra cost by two main arguments. First they argue that the expenditure on an earth road which is not properly constructed may be lost altogether, since it can rapidly become irreparable under the combined impact of rain and traffic, and will have to be abandoned in favour of an entirely new trace. This argument seems to me to ignore the time factor. A road built to the proper standards of the professional engineer is bound to require a longer period for preliminary planning and surveying. This is implicit in the engineers' argument, which ascribes the defects of the amateurs' efforts to insufficient attention to the crucial factors of drainage and alignment. But if the road is likely to open up land for production, and the resulting annual net output (as I began by suggesting) considerably exceeds the capital cost, then the loss through delay may likewise considerably exceed the loss which may arise through the necessity of subsequently reconstructing the less durable type of road, built by amateurs in a shorter time.[1] Secondly, it is never possible to be certain, in an unplanned economy, that the invitation to develop the adjacent land, which a road offers, will in fact be accepted. There were certainly wide differences between the response of different districts in British North Borneo. This makes it desirable to spread the funds available over as many projects as possible, and the number will be reduced if professional standards must always be observed.

---

[1] *The Economics of Under-developed Countries*, p. 151.

[2] It would be more correct to describe the excess stocks as ' disguised loanable funds ' than as ' disguised savings '.

The second of the engineers' arguments concerns the type of vehicle which the road will enable to be used. The earth roads built by the administration in British North Borneo were known as 'jeep tracks', because they could only be negotiated by vehicles of the jeep or Land Rover type, having a four-wheel drive. The professional engineers claimed that earth roads built to their standards would take a 3-ton lorry or a small bus. The greater capacity of these vehicles would reduce the cost of carrying a ton or a passenger a mile by more than enough to offset the higher capital cost of the road itself. From various enquiries it appeared that cost per capacity ton-mile would be reduced from about 1s. 5d. by Land Rover to about 8d. by 3-ton lorry; and these figures indicated that the extra capital cost of the road built to professional standards would be justified at a very low level of traffic. Nevertheless this could not be regarded as conclusive. In the first place, the use of the bigger vehicle would tend to diminish the frequency of service, if existing utilisation of capacity were to be maintained; or, if the public or public obligations laid on operators required a service of given frequency, it would diminish the utilisation of capacity. Secondly, the larger vehicles are specialised for carrying either goods or passengers; Land Rovers carry both. This adaptability to the traffic offering, especially in conditions where it is sparse and unpredictable, may enable the latter to achieve a better utilisation of capacity. Finally, it is observable that many vehicles of the Land Rover type are still in use in areas served by roads which have been metalled and sealed and thus brought to a standard superior to that of the professionally built earth road. Thus the apparent possibility of achieving the economies of larger vehicles is often not exploited in practice. The reason for this is that in a developing agricultural country there will always be roughly constructed private farm or estate roads running off the public system, and negotiable only by small vehicles of the Land Rover type. At the point of contact with the public system, the farmer could remove his goods into the larger and more economical vehicles which professionally-built public roads will carry. But the cost or inconvenience of removing the goods will deter him from this unless the economies of the larger vehicle are substantial. These considerations do not refute the engineers, if they mean that larger vehicles *can* offer economies. The question is whether what can be will be.

---

[1] I have given reasons for being suspicious of claims that a road will open up land, unless the other factors of production are available. If they are not, my point here has less weight. But then the question is not whether the road should be a professional job or an amateur one, but whether it should be constructed at all.

The above arguments concerning low cost roads have not considered at all the influence of the standard of the road on the running costs of vehicles of any given type. Remakably little seems to be known about this, although it is clearly important as a factor in deciding at what level of traffic it is justifiable to raise the standard of a road. The colonial division of the Road Research Laboratory is interested in this question, however, and evidence on it will bring joy to many road planners working in the dark.

It would be rash to draw any general conclusions from limited experience of a country as small as British North Borneo. But of one thing I am certain. I have said that when roads are under discussion, the views of engineers and economists are apt to collide. This does not mean that they should steer clear of each other. At present the economists are ignorant of the engineering problems, and the engineers are prone to over-simplify the economic ones. Anything which contributes to the exchange of ideas and knowledge between the two will be of great benefit to the practice of road planning.

Worcester College,
Oxford.

ROAD TRANSPORT—A GAP IN NATIONAL STATISTICS

By D. L. Munby

Hitherto, transport statistics have largely survived as a residue of Victorian and Edwardian reformism. In many other fields attempts have been made to produce twentieth-century statistics, but not, up till now, in the field of transport. Recent sample surveys of road haulage and of the road system, conducted by the Ministry of Transport and the Road Research Laboratory respectively, have used modern techniques to fill gaps and explore important fields of enquiry. But a lack of money seems to have hampered these new endeavours. Thus only a little over £3,000 was spent on the 1958 Road Transport Survey [6, p. 6].[1] At the same time existing statistics have not been consistently overhauled so as to produce a comprehensive picture of what we do know. There seems, however, to be a change taking place. In inaugurating work on the new Blackwall Tunnel, the Minister of Transport referred to the need for more information about traffic in London, and for examination of the statistical basis of suggested improvements. There are signs that, within the Ministry, statistics are to be taken more seriously. The time is therefore perhaps ripe for making suggestions about what is needed in this field, so that policy can be firmly based.

I. The Stock of Vehicles

The main source of information on the stock of vehicles in the country is registration returns [1]. These cover all types of vehicles with licences current at any time during the quarter ended September 30th each year—that is, all vehicles other than Service vehicles and vehicles exempt from duty because not used on public roads. They also give new registrations for the year ended December 31st.

1. *Goods Vehicles*

a. Analysis of the Stock. For goods vehicles a second source of information is the reports of the Licensing Authorities [2], which give figures of the stock of vehicles at the end of the year. These now include all goods vehicles, whether coming under A, Contract A, Special A, B, or C licences. But from 1948 to 1952 they excluded British Transport Commission vehicles. The following are some of the inadequacies of the figures.

(1) The discrepancy in the dates (September 30th and 31st December) makes it impossible accurately to combine the two sets of data. If September 30th is a necessary date for the registration figures, there seems little reason why the Licensing Authorities should not publish their figures for the same

---

[1] Reference is made to the chief sources of information on road transport by number, the items being detailed in the Appendix.

date. (It is indeed possible at some trouble to make some reconciliation by using the monthly new registration data from [1], though estimates have to be made of the numbers of vehicles ceasing to be registered.)

(2) In both sources, vehicles are classified by unladen weight, whereas the normal international classification is in terms of carrying capacity (as with E.C.E. [15]). The British Transport Commission [4] publishes its statistics in these terms, but does not publish them in terms of unladen weight. Unladen weight figures for B.T.C. vehicles are included in the Licensing Authorities' figures, and are separated out by the Ministry of Transport and published in the *Annual Abstract of Statistics* [3] (but nowhere else). For the missing years 1948–1952, when they are not covered by the Licensing Authorities' figures, only estimates are possible.

There is no consistent relationship between a given unladen weight and a given capacity. The 1952 and 1958 enquiries [5, 6] asked questions about this, and published some figures (for some groups) comparing the two measures ([5], Table 3; [6], Table 2). Apart from these and the British Transport Commission figures, we have no information about the carrying capacity of British goods vehicles. This gap obviously needs remedying, even if carrying capacity is a less precise concept than unladen weight, and some rather arbitrary measure has to be chosen (e.g. manufacturers' rather than users' figures).

(3) The registration figures classify vehicles also by type of fuel used, which is not done in the other source. This is of great importance, as the use of diesel fuel is growing at the expense of petrol. The 1952 and 1958 enquiries published information about the use of diesel fuel according to carriers' licences as well as unladen weight, and the results [5, Table 4; 6, Tables 3–4] suggest that diesel vehicles have (or may have) a different pattern of usage from petrol vehicles. These figures should be available every year.

(4) The registration figures, which give great and valuable detail in several directions (e.g. in breaking down the figures between England, Wales and Scotland, in distinguishing farmers' and general goods vehicles, etc.), have been rationalised recently and presented more clearly. Rather less detail than formerly is now available, as various special vehicles (showmen's goods vehicles, tower wagons and local authorities' watering vehicles) are no longer (since 1956) broken down by size categories.[1] This is perhaps not so important, as ' general goods ' will in any case include many other types of specialised vehicles. It might, however, be worth considering whether a census of all sorts of special vehicles might not be taken at intervals.

To summarise, what is required for goods vehicles is one comprehensive set of figures showing the annual stock which will be comparable in coverage and time with the new registration figures, and will include data on unladen weights, carrying capacity, fuel used and type of carriers' licence. By rationalising the existing figures in this way, they would become much more useful and easy to interpret.

---

[1] In the detailed breakdown of the registration figures, moreover, there is no consistency in definition of the open-ended group of heavy vehicles: the line is sometimes drawn at ' 9 tons and over ', sometimes at ' 6 tons and over '. And practice differs from year to year, according to no discernible principle. As the development of the larger vehicles is one of the most important recent changes and one that needs to be studied from various points of view, the failure of the detailed figures at this point is particularly reprehensible.

b. Industrial Ownership of Vehicles. Practically nothing is known about the ownership of vehicles of different kinds. The only detailed information we have is of the stock of vehicles owned by farmers (from [1]) and by the British Transport Commission. The 1958 survey [7, Table 7], gives the distributon of C Licence vehicles by the Orders of the Standard Industrial Classification.

Capital investment in vehicles in post-war years has been a by no means negligible part of total capital formation. Overall figures are available in the National Income Blue Books, and detailed figures of yearly investment are obtained in the Censuses of Production [8]. These give purchases of new and second-hand vehicles and disposals by industry groups, so that it is possible to see which industries are investing in vehicles.

There are, however, three gaps in these figures for our purposes:

(1) The figures are in pounds sterling, and can be related to output figures only in general terms.

(2) We have no information about the stock of vehicles held by different industries (or the value of the capital stock) for any date.

(3) Vehicles include passenger cars as well as goods vehicles. Any attempt to get a true economic picture of the transport industry requires that these deficiencies should be made good. Nor would the required information be in the least difficult to collect, or involve industry in any difficult calculations. If the Censuses of Production and Distribution asked for the numbers of vehicles bought as well as the cost, and distinguished between passenger cars and goods vehicles, a large gap in the national figures would be filled. Stock figures could perhaps be collected at intervals through the Censuses of Production and Distribution. There ought, in any case, to be no difficulty in collecting the information through the Licensing Authorities, in the case of Contract A, B and C licences. It ought not to be impossible to find out the industrial category to which they belong for Census purposes, and the Licensing Authorities could publish this information with their other figures.

2. *Cars*

Some of the general remarks about the registration returns apply also to cars and motor-cycles. An additional difficulty arises from the fact that some cars are classified by horse-power and some by cylinder capacity, the former representing the system of taxation up to 1947. While these cars still exist, it would be worth while publishing a comprehensive classification in terms of cyclinder capacity, even if it meant taking rough 1947 conversion factors.

The major problem in the case of cars is our complete ignorance as to the number of cars used for business purposes, and the number of new cars bought by industry. The estimates made by the Central Statistical Office are admittedly little better than guesswork. From 1946 to 1951, about 50 per cent of the value of new cars was taken as consumers' expenditure rather than business outlays, and after 1951 it has been assumed that business outlays on cars have varied roughly in proportion to total fixed capital formation in industry.[1] The present situation is somewhat ludicrous in view of the importance of the question, and of the extent to which estimates of the return on investment in roads depend on the valuation of time saved and its distribution between business and pleasure. Much of this could easily be remedied, if we have the data suggested above from the Censuses of Production and Distribution.

For private cars, we would like to know more about ownership by income groups. Information is available about this in the Savings Surveys of the Institute of Statistics and the C.S.O., and this seems the best way to collect what is required. We need a continuous survey of ownership and new purchases of private cars.

II. A Census of Road Goods Transport

The lack of a Census of Road Haulage is one of the biggest gaps in the whole set of national statistics. Whether one looks at the matter from the point of view of the need for further information about transport (to aid policy in the transport field), or from the point of view of inter-industry relations (the need for a more comprehensive input-output table for the economy as a whole), the gap stands out.

The reason usually given for the failure to provide this simple tool of rational policy is the small scale of the road haulage industry. With 1,308,000 vehicles with carriers' licences in the hands of over half a million operators in June 1959 ([7], p. 107), it is argued that there are too many returns to be filled in, and that the operators would be unable to provide detailed information, because they do not keep detailed records. But it would seem that this argument is often the perfectionist's argument for doing nothing, whereas in practice a little information would be much preferable to nothing. The argument from mere size is not convincing.[2] The first Census of Distribution in 1950 covered over half a million shops. Like the Census of Distribution, a Census of Road Haulage would not need to be carried out at more than given intervals. But it would provide the indispensable framework for sample inquiries in other years.

Before turning to what would be required from a Census of Road Goods Transport, we must look at the two existing sources of information: the 1952 and 1958 enquiries and the Censuses of Production and Distribution.

The 1952 *and* 1958 *enquiries* [5, 6, 7] were pioneering sample surveys in the field of road haulage. It is something of a scandal that Section 16 of the 1933 Act (now Section 186 of the Road Traffic Act, 1960) was never used until 1952 to collect information about the operations of road hauliers, though the Act lays down that holders of carrier's licences 'shall keep . . . current records showing . . . particulars of every journey of a vehicle on which goods are carried under the licence, showing the greatest weight of goods carried by the vehicle at any one time during the period to which the record relates and giving the description and destination of the goods carried'. (1960 phrasing, which is slightly different from that in the 1933 Act. The duty is subject to regulations and exemptions).

---

[1] *National Income Statistics—Sources and Methods* (Central Statistical Office, 1956), pp. 124–5.

[2] In any case, most operators are C licence holders. At the end of 1958 there were 547,000 of these to 34,000 B licence and 28,000 A licence holders.

The 1952 and 1958 enquiries are unfortunately very limited, the latter enquiry being not so thorough as the former. In effect, hard information is available only for two particular weeks in September 1952 and April 1958, and, as we know nothing from other sources about either seasonal patterns or trend, any conclusions from their results are highly dubious, if attempts are made to derive information from them about the increase in particular types of business.

They have been used, in conjunction with estimates of vehicle-mileage from the Road Research Laboratory sample census points [11], to estimate the total volume of goods carried by road in 1952 and 1958, in terms of tons and ton-miles, as well as to produce a monthly index of the volume of road haulage. [5, 6, 7, 16]. Unfortunately, we have no information about the seasonal pattern of road haulage in terms of tons and ton-miles. If we assume that the figures of vehicle mileage are correct, there seems every reason to suppose that there will be seasonal fluctuations in the relationship between vehicle mileage and tons and ton mileage. There is a great range of variation in this relationship according to size of vehicle, and kind of licence [6, Table 14]; one would expect it also to vary seasonally.[1]

In detail, the two enquiries are not comparable, nor did the 1958 publication clearly present the two sources side by side. But a more serious question is the validity of the use of the ton and ton-mile as a measure of road haulage activity, whether for the road haulage industry or in comparison with other forms of transport; to this matter we return later.

The Censuses of Production and Distribution [8, 9] give some information about transport payments, but, as this was collected with other aims in view, it is inadequate and incomplete.

Before the war, no information was collected in the Censuses of Production about transport payments by firms. Thus the gross output of a firm included payments made for the outward transport of goods from factory to seller, in so far as these were included in the sales price of the goods.

Every seller has a choice between selling at an ex-works or a delivered price; in the latter case, he makes himself responsible for the transport payments involved in delivery of his goods. But he may equally be responsible for these, even if he quotes an ex-works price and adds the delivery charges to his bill. (Responsibility for payment of delivery charges in the first place is no evidence of an effective delivered price system).

---

[1] The arguments in [7, pp. 117–8] referring to changes in vehicles' load factors seem to ignore the fact that there are two senses in which the term 'load factor' can be used and the peculiarity of the particular use in the available source.

Transport costs can be met in three ways: the seller may use his own transport, the buyer may do the same, or there may be an outside contractor (for Census of Production purposes a firm's ' separate transport organisation ' is equivalent to an outside contractor).

In the Censuses of Production for 1948, 1949, and 1950 firms were asked to state the transport payments ' made to other firms or credited to firms' own separate transport organisations for transport outwards on finished goods sold '. Thus it is possible for these years to deduct from sales the amount of outward transport cost included in sales price, in so far as outside contractors were used or firms had separate transport organisations, but not in so far as firms did not have a separate transport organisation. With this latter exception, ' gross output ' excluded the outwards transport element in sales price; ' net output ' would equally be clear of this extraneous element in so far as the gross cost of materials (including inwards transport costs) is deducted from gross output. Net output thus represents the value added by the firm in its own proper sphere of operation, with the addition of such ' transport output ' as did not result from a separate transport organisation.

For 1951 and later years a different method was adopted, whereby firms were asked to state their total transport payments ' paid (or credited) to other firms, and undertakings, and to their own separate transport organisations, for transport during the year, whether for outwards transport of finished goods sold or inwards transport of materials and fuel purchased '. A more complete picture of the total transport element in sales price was thus obtained, but at the expense of a change in the valuation of gross output (back to the pre-war usage). As there is no distinction in the figures between transport inwards and outwards, it is necessary to define gross output to include payments for transport outwards, so that gross output is equal to sales made, adjusted for changes in stocks and work in progress. But there is no change in the case of net output, for which total transport payments have to be deducted.

The Census of Production excludes agriculture, transport, distribution, and the service sectors, as well as separate transport organisations of industries covered by the Census. In so far as transport payments are made by the Census of Production firms in their dealings with other sectors not covered by the Census, they will be included in the totals. But it will be to some extent accidental whether this occurs or not. Thus firms may have bought raw materials from the agricultural sector at delivered prices, and may have sold to wholesalers or retailers at ex-works prices, leaving the transport charges to be made by the purchasers. It is impossible to estimate from the figures available how far in fact this occurs or not. The National Income Blue Book gives figures of the contribution of the Transport and Communications sector to the Gross Domestic Product, but these represent the ' net output ' of the sector, whereas it is the ' gross output ', which represents the total amount paid to it by other sectors, in which we are interested. If we knew the gross output, we could compare it with the total payments recorded in the Census of Production.

Some figures are available from the Census of Distribution as to transport payments made by wholesalers in 1950, but this again covers only a part of the field. For the agricultural sector, estimates are made yearly of the changes in costs of transport and marketing in the Annual Price Reviews, but no estimate of the gross sum. It is thus impossible to obtain any figure of total gross output of transport from official sources. Figures are, of course, available for the railways, but nothing is known of the total of payments for road transport services. This is the fundamental figure we want to know. The rough estimates that have been made by Rudd and Walters[1] are better than nothing, but they only give us figures for 1949, 1950 and 1953.

How can these gaps be filled in by a Census of Transport?

(1) The first information that a Census would ask for would be the number of operators, their vehicle fleets (analysed by unladen weight, carrying capacity, type of fuel and type of licence), and their employees. It would then be possible to get a picture of the industry in terms of size, such as we do not have at the moment.

(2) The second basic piece of information would be about total costs and receipts for each enterprise. As this information must be available for income tax purposes, it would create no difficulties. Most firms (at least those of any size) would equally find little difficulty in providing the relatively simple information about costs that is required. The main items required are fuel costs (distinguishing petrol and diesel if possible), direct labour costs, repairs and depreciation, and licence duties and overheads. British Road Services (B.T.C. Report, 1959, Vol. II, Tables VI–3, 3a–h) give rather more detail, but even this should not be beyond the powers at least of the larger firms.

For receipts, what we want is a breakdown on an industrial basis. This means that customers must be classified on the basis of the Standard Industrial Classification (which ought not to create major difficulties, if help were given by the Census authorities). No doubt many small firms will not be able to give any information on this matter, if their customers are many and none large. But many hauliers must do considerable business with a relatively small number of firms, and information about these large customers would be extremely valuable. It would be better to have information about customers with bills of more than £x00 rather than no information at all.

This information is more important than data about classes of goods carried or tons and ton-miles, though if the latter could be collected as well so much the better. (See later).

If this information could be obtained for individual vehicles, one would begin to have a picture of how the assets of the industry are really used and of the spread of costs for different kinds of operations.

(3) We want operational statistics. The most basic figures of vehicle-

---

[1] E. Rudd, 'Estimates of Expenditure on Road Transport in Great Britain', *Journal of the Royal Statistical Society*, Series A, 1952, Part II, pp. 179–192. A. A. Walters in G. Walker, 'Competition in Transport as an Instrument of Policy', *Economic Journal*, September, 1956, p. 413.

miles should not be difficult to collect. (Again, if these could be made available for individual vehicles, we should be very far ahead, and this goes for what follows.) We also want to know what proportion of these vehicle-miles are loaded, and what proportion empty or half-loaded. No doubt some very rough approximations are needed to measure ' half-loads ', as was done in the 1958 enquiry [6, paras. 35–41], but the high proportion of ' end to end ' work suggests both that it is not so difficult to collect figures on loadings as might be thought the case, and that errors may not be all that important.

We do not want only vehicle-miles; we also want to know on what routes vehicles operate. Though the collection of information on this may prove difficult, it should not be impossible. Where vehicles engage in ' tramping ' within a small area, detailed information is not required. Where vehicles go long distances, or are used on a regular ' middle-distance ' route, information would be easiest to obtain and that which is most needed.

Again, we need to be able to tie in this information, if at all possible, with the information on an industrial basis. Ideally, we should like to know what vehicle-mileage, and on what routes, is for different industries.

(4) We want to have information about purchase and sales of vehicles and other capital assets, of the same kind as is collected for the Census of Production. (This information has been collected from companies for 1956 and successive years for net capital expenditure.[1])

It is accepted that some of this information could not be collected from small firms, and a simplified form would be required for firms with 1 or 2 vehicles. These firms are indeed of special importance in the road haulage industry. But once their general characteristics were made clear in a full census, it would be possible to sample them in subsequent years.

C licences require special treatment. Most of their owners already complete the Census of Production or Distribution; but little, if any, of the relevant information has been collected through these means. All that would be required would be that the relevant figures were collected with these censuses. Some information about them is available from the 1958 sample enquiry [7, p. 115]. 40 per cent of vehicles are engaged in distribution, 16 per cent in construction, and 9 per cent in the food, drink and tobacco trades. In these cases we want to know what kind of work is involved more than some other details. In the case of shops with delivery vehicles, for example, it is more important to know about the kind of service provided and the radius served than details of costs and receipts, or even vehicle-miles. In the case of building, it is more important to know whether the vehicle is merely engaged on work from depot to building site (or in collection of building materials from quarries, brickworks, etc.) than to have other kinds of information. Similarly, for food firms which deliver from their factories to shops, the radius of operation may be more important than any other information.

The same applies to farmers' vehicles, about which separate information is available in the registration returns [1]. This is the only industry about which we have any information about its stock of vehicles, and this is a mere accident of the fact that farmers have been specially treated for licence purposes. It is perhaps more important here again to know the special uses to which these vehicles are put than some of the more general matters which we need to know about ' transport for hire and reward '.

---

[1] See *Board of Trade Journal*, February 6th, 1959.

In short, for some specialised forms of C licence, particular questions will have to be asked. Nor will it be necessary to ask all the questions that are asked of the haulier plying for hire, though some special questions may be asked, as for example whether it is mainly raw materials or finished products that provide the main use of an industry's vehicles.

III. The Road System

A. Roads.

1. Up to 1955–6 the main source of information about roads was to be found in the Road Fund Reports [10]. In that year the Road Fund was abolished, and responsibility for Scottish roads was handed over to the Scottish Home Department. As a result, there is now no single publication (except for the *Annual Abstract of Statistics*) which gives information about the total road system; English and Scottish figures are published in two different reports [11, 12].

Complications arise from the fact that the Scottish local authority year ends on May 15th, in contrast with the English year's end on March 31st. For most purposes the difference may be very small and insignificant, but it can lead to confusion.[1]

The published figures give details of road mileage for different administrative classes of road for different types of local authority in England, Wales and Scotland. For the same subdivisions, expenditure is also given broken down into maintenance and minor improvement, new construction and major improvement, cleansing, administration and so on. These figures are also divided according to the source of funds (local and central), and similar figures are available on the basis of grants paid by the central government, which differ in timing from the figures of approved expenditure.[2]

2. Sample surveys conducted by the Road Research Laboratory have revealed many unknown characteristics of the road system (such as data about the number of bends, gradients, lane markings, footpaths, street lighting, Zebra crossings, width of roads, etc.). Unfortunately, this information is scattered throughout a number of technical papers and in the annual reports on Road Research [13, 14], and is nowhere brought together, so that it is possible to have a clear picture of the road system and what we know about it.

---

[1] Thus the *Annual Abstract of Statistics* (No. 96 (1959), Table 227) gives figures allegedly for the year ended March 31st, for Great Britain, whereas it seems clear from Scottish sources that these are merely additions of the Scottish and English local authority figures for the years since the changeover.

[2] Up to 1951 there are no figures of actual expenditure on (as distinct from grants for) trunk roads.

The main deficiencies of our knowledge in relation to our concern about the use of the roads are as follows:

(1) Data are available from the Road Research Laboratory as to the mileage of road in built-up and non-built-up areas, but these are not regularly published with the other figures. This should be done.

(2) There should be some publication where the main known characteristics of the road system are described.

(3) We want to know much more about the variations in expenditure on roads not merely between different types of authority, but as between different authorities of the same type. In other fields of local government expenditure, much information is available on these lines, e.g. in the case of housing, rates, and education, either in official publications or in publications of the Institute of Municipal Treasurers and Accountants.[1]

(4) The items of expenditure in the normal returns are not, in fact, a complete return of all expenditure incurred on the roads. Expenditure on street lighting is not included (it would have added some £18 m. to the total of £109 m. in England and Wales in 1956–7; [3 (1959), Tables 304–5].

More important, we have no estimates of the likely police cost of the roads. It would be highly useful if some calculation could be made of the cost of signalling on the roads.

B. The Use of the Roads.

Traffic counts have a long history, and various more or less comprehensive traffic censuses have been carried out. Up to the 1954 Census these were taken at 5,830 selected ' permanent census points ', and, while providing information about these specific important points in the road system, the results were so biassed as to preclude any calculation of total traffic on British roads. In 1956, a new kind of census was initiated, taking 100 points in England and Wales selected by some allegedly scientific procedure, from which it is thought that the volume of vehicle-mileage can be properly estimated. This is an annual census for a week in August. It is supplemented by a continuous count at 50 points throughout Great Britain. The census covers only Trunk and Class I roads, whereas the continuous census also includes Class II and III roads (only unclassified roads are excluded). From these have been built up a series of estimates of vehicle-mileage. For the annual census these are published for the nine engineering divisions in England and Wales, according to different categories of vehicles, differentiating between workdays, Saturdays and Sundays, and also between urban and rural roads. Similar results are available from the 50-point census from which monthly data are also reported [11, 16], but these figures are not broken down on a divisional basis.

The method used in these censuses has been discussed in an article in *Applied Statistics* [17], as well as in [7, 11 and 16]. Their main use is to produce an estimate of the total number of vehicle-miles travelled, and their changes through time. While the results seem reasonably convincing, the method does not at first sight inspire confidence. There are some 95,000 miles of road (excluding unclassified roads) in Great Britain, and 50 points can hardly represent an adequate sample of the road system. There are some 21,000 miles of Trunk and Class 1 roads in England and Wales, and even 100 points seem a rather limited sample. When divided into nine divisions and 4 classes of road (Trunk and Class 1, each differentiated according to rural and urban roads), there are 12 sectors which have only one point, and six with only two. It is admitted that not too much reliance should be placed on these results in detail. [11(1956–7), pp. 25–6].

---

[1] See Interdepartmental Committee on Social and Economic Research, *Guides to Official Sources, No. 3: Local Government Statistics* (H.M.S.O., 1953).

Again, the division of the statistics between England and Wales and Scotland produces difficulties. Annual censuses are conducted separately in Scotland, but their results are not published.

Further detailed enquiries have been made by the Road Research Laboratory [13, 17].[1] One of these involved the study of four areas (Buckinghamshire, Dorset, Lancashire (other than county boroughs), and Leicester C.B.). Fifty points were chosen on all types of road (including unclassified roads) in each area, and from this some idea of the range of traffic density on different parts of the road system can be estimated.

Overall estimates have their uses, but we need to know far more about the degree of under-utilisation and overcrowding of different kinds of road. It is important to be able to compare the use of the roads with variations in expenditure on their maintenance, so that we can form judgements as to whether too much is spent (or too little) on particular kinds of roads, or in particular areas.

We also need to know something about the flow of traffic throughout the country. It ought to be possible to produce maps of traffic flow showing the main streams at the very least.

It does not seem that the information at present available is adequate to help in the formation of a coherent and rational policy.

IV. SPECIAL PROBLEMS

1. *Ton-Miles and Vehicle-miles.*

The 1952 and 1958 sample enquiries were fundamentally concerned to discover data from road hauliers in the form of tons and ton-miles. Goods were classified according to the Standard International Trade Classification with some modifications. Though useful for many purposes, this clashes with that of the Standard Industrial Classification, and therefore prevents any use being made of the figures to study the pattern of inter-industry relations and the way different industries generate demand for transport.

Fundamentally, these surveys were concerned to measure the volume of road goods traffic in terms of tons and ton-miles, so that comparison could be made with other forms of transport, notably railways. But the question needs to be raised as to whether the ton and ton-mile are acceptable as measures of volume in transport.[1]

---

[1] I omit reference to detailed local studies of traffic by origin and destination that have been carried out (e.g. in Oxford, or for the London-Birmingham motorway).

It is much more difficult to obtain figures of tonnage and ton-mileage from road hauliers than it would be to obtain figures of vehicle-mileage. The vehicle is the unit of the operator, and figures of tons are not operationally significant to them or normally collected by them (at least in the case of small firms). This surely raises the question as to whether these figures really matter, except for the fact that convention has always kept transport figures in this form.

An index of volume of production may be wanted for various purposes. One is to measure changes in the employment of factors, another to measure the productivity of factors, a third to measure changes in demand. The simple model, which is often in people's minds when discussing these matters, is one which assumes that there are three variables (price of output, quantity of output and quantity of factors).[2] Between these three variables, some simple relationships are assumed, which produce measures of change in output and productivity. This model may be convenient for many manufacturing industries, where factors may be broken down into labour, raw materials and capital, the former two varying more or less with output, and the complications which arise in the case of capital being treated separately. But some of the complications, which arise elsewhere with capital only, confuse the issue generally in transport. Here the problem of utilisation of capacity arises at the stage of variable inputs, and has to be treated as a separate item, if we want to analyse what is happening.

(*a*) If we are interested in changes in ' productivity ', we find very simply that a change in final demand (in terms of tons and ton-miles) may lead to very small changes in inputs (e.g. a little more petrol and vehicle wear-and-tear), if it leads to greater utilisation of vehicles. An increase in productivity can come about, not because of a more efficient use of factors or a greater ' effort ' per unit-factor, but because demand is now more adjusted to the unit of supply. It is important to be able to distinguish these two forces, improved utilisation of equipment and lower costs per unit of supply. Concentration on tons and ton-miles confuses the two forces, which can in fact be easily measured in transport in terms of vehicles and their cost/receipts, on the one hand, and, on the other, the proportion of their capacity that is loaded, or of empty journeys.

No doubt the ultimate demand of the consumer is for a given consignment to be carried a certain distance, and, if the ton is the proper measure of the volume of his demand, ton-miles will measure the volume of consumers' demand. We are also interested in knowing how, for the economy as a whole, demands from consumers lead to demands for factors. It is on this relationship that the overall productivity of the economy depends. The more the organisation of the road haulage industry is adapted to meet the demands of consumers as they arise with the smallest number of vehicles, the richer the community will be. This may be a more important element in the productivity of the industry than any reduction in running costs or improvement in design of lorries. But it still remains true that there is an advantage in being able to distinguish these two separate influences in the case of road haulage.

---

[1] This is discussed by Glover in [7].

[2] Also no doubt price of factors; but this complication does not concern us here.

(*b*) It is not clear that the ton is of great significance as a measure of volume in transport. The quality differences between tons are so great both on the demand and on the supply side as to rob the measure of much value. It would seem that these differences are greater than those that arise in the case of most other industries. But here again it may be that the trouble has arisen partly at least from the choice of the measuring rod.

From the point of view of the transport operator, roughly three characteristics of commodities have to be considered, weight, cubic capacity and loadability. It is a combination of these three characteristics which determines whether a given lorry of say 6 tons capacity will be entirely filled or half-empty with a given consignment of 3 tons. If a given consignment of 3 tons weight has to be carried in a lorry of 6 ton capacity, there is little point in producing statistics which show that the lorry is half-full on such journeys. Both from the supply side and the demand side, the whole lorry is the relevant unit that is bought and sold. We do, on the other hand, want to know whether lorries are half-full or completely full in the ordinary operational sense in which the question would be answered by the haulier.

It is, however, the meaningless sense of a ' half-loaded ' lorry (according to which this merely refers to the fact that, if there were a different load weighing the same as the actual load which has to be carried, it would be possible to carry twice as much in weight in the given lorry), which is used in the sample enquiries. Thus Table 7 of the 1958 enquiry giving the proportions of vehicle mileage run loaded and empty refers merely to weight, and, as is noted, ' light but bulky goods may fill all the available space in a vehicle without filling it to its weight capacity '. (para. 33). Having decided to measure ' intensity of use ' (para. 61) by ton-mileage, as well as vehicle mileage (which is more significant), it is not surprising that disclaimers have to be made to obviate the misleading political use that could be made of these figures from ' C ' licence operators (para. 68–69). What we do not have from this enquiry is any idea of the extent to which vehicles were run loaded or half loaded in terms of the everyday judgment of the operator.

It is for these reasons that it has been suggested earlier that figures should be collected on a vehicle basis. Vehicles can easily be classified by their ' tonnage capacity '; the ton measure may be more useful for a vehicle than for a load (with the obvious exception of some specialised vehicles). If we knew the costs of operation of a vehicle of e.g. 5 tons capacity, the receipts it obtained in a year, its annual mileage, the loaded, half-loaded and empty mileage (on our standard), and the customers (by industry) it served, we would know a great deal more of use than can be obtained from figures of tons and ton-mileage. Thus the difference between different industries in payments per loaded vehicle-mile would be far more significant than differences in payments per ton.

No doubt this has implications also for railway statistics, where the wagon may be a more significant unit than the ton.

2. *Prices and Costs.*

We would like to know something about prices and costs in road haulage. Cost information should become available from a proper census type of enquiry such as we have outlined. Price information is much more difficult, and yet information on this point would help us to understand what has been happening in the changes in traffic as between rail and road, etc.

From a census of a full type, it ought to be possible to discover the prices paid per vehicle-mile for different kinds of industries, as we have suggested earlier, and different vehicle sizes. From this information we could deduce some kind of price index or set of indices. It is quite likely that this would be as meaningful and useful as any attempt to collect prices in terms of tons and ton-miles, which is where temptation leads us at first. Clearly, for tons of coal and such-like commodities, payments per ton and ton-mile are reasonably meaningful. But, for the reasons given earlier, this is hardly true of the more highly fabricated commodities.

It is when we come to compare road and rail that difficulties about price arise. No doubt the degree of substitution between different forms of transport has been exaggerated, but there must be a range of commodities for which it is important. Here the customer will no doubt think in terms of the charge per ton-mile. It is only in so far as the wagon is comparable to the lorry that payments per vehicle-mile can be related to railway statistics. The problem is likely to remain permanently a difficult one, with no clear-cut solution. Perhaps, meantime, the government could collect figures from the Road Haulage Association, which would form the basis of a rough index number.

3. *Road Passenger Statistics.*

In the passenger field, as a result of the fact that all bus companies make fairly comprehensive returns, we have much more detailed statistics than in the haulage field. The main publication is clearly produced and reasonably complete [18]. We have information about journeys, receipts, vehicle-miles and fares. The main gap is any information about passenger-miles for buses other than those operated by the British Transport Commission, information which would be very useful, even in the form of rough estimates.

4. *Business Cars.*

From what has been outlined above, it would be possible to obtain a much clearer picture of the movement of goods and persons by ‘ public operators ’. We still need to know much more about the use of private cars for business purposes. Once we have information about the stock of such cars, information about their use would not be difficult to obtain on a sample basis. Figures of mileage and petrol consumption would be the chief requirement. (The latter has already been collected in some censuses of production for all vehicles.) Further broad information about categories of use would also be fairly easy to obtain, and would help to shed light on our present darkness.

5. *Other Sources.*

The above summary excludes statistics about road accidents and road offences, which are not directly germane to our concerns. It makes no attempt to summarise what is available from the past; this has been done up to 1950 by F. A. A. Menzler.[1] Nor does it attempt to cover the sources which are not primary (such as *Basic Road Statistics*, published by the British Road Federation, and *The Motor Industry of Great Britain*, published by the Society of Motor Manufacturers and Traders), or entirely derived from industry (such as data published in *Commercial Motor*). It has confined itself to a survey of the original sources of information made available by the Government.

V. CONCLUSIONS

Existing information available to the government in formulating policy is clearly inadequate and badly presented. There are signs that improvements may occur. It may be worthwhile giving a list of priorities for improvement:

(1) The figures need tidying up and presenting clearly. The compilation of an annual digest of road statistics (perhaps of transport statistics as a whole) would make the existing information of considerably greater utility. It is a sad comment on government parsimony that the only comprehensive source of road statistics has been *Basic Road Statistics*, published by the British Road Federation, an ex-parte body, which merely republishes government figures with all their deficiencies plus a few more of its own.

(2) A summary of the characteristics of the road system as revealed by the researches of the Road Research Laboratory, bringing together the statistical results of their many enquiries, would be of great value. It is at the moment extremely difficult easily to discover the qualitative characteristics of British roads (e.g. width, crossings, etc.).

(3) We need more details of expenditure on roads, both to produce a more meaningful aggregate figure and to enable us to compare variations in expenditure in different parts of the country.

(4) The same applies to traffic flows.

(5) We must have more information about business ownership of cars and lorries and their use.

---

[1] 'Road and Rail Statistics' (*Journal of the Royal Statistical Society*, Series A (General), Vol. CXIII, Part III, 1950). Reprinted in M. G. Kendall (ed.), *The Sources and Nature of the Statistics of the United Kingdom.*

(6) Fundamentally, we need a figure of total expenditure on road haulage. A Census of Road Goods Transport would fill in one of the major gaps in our knowledge of the flow of expenditure throughout the economy. Much of the valuable information we could obtain from it could perhaps be obtained by continuous surveys of the kind conducted in 1952 and 1958. Part of the issue depends on whether figures of tons and ton-miles are more useful than figures of expenditure. If the latter is in the last resort what we want, it must come from a Census, even if only to provide a base for future sample enquiries.

A little investment in statistics would bring in large returns, if the aim of policy is to spend money intelligently on the roads. It would not seem that the suggestions made above would be either wildly expensive or inherently difficult and/or time-consuming for busy people. At least the matter merits consideration.

Nuffield College.

APPENDIX

1. *Road Motor Vehicles* (Ministry of Transport). Annual. Includes Annual Census of Vehicles, New Registrations and Taxation Receipts. Monthly returns of new registrations, and quarterly returns in more detail, are available from the Ministry of Transport. Quarterly statistics were printed formerly under the heading *Mechanically Propelled Road Vehicles*, but this printed series ceased at the end of 1956.
2. *Summary of Annual Reports of the Licensing Authorities* (Ministry of Transport). Annual. Gives details of applications for licences and statistics of the quasi-judicial proceedings of the Licensing Authorities, also details of the vehicles and operators in the different licence categories.
3. *Annual Abstract of Statistics* (Central Statistical Office).
4. *British Transport Commission: Annual Report and Accounts.* (House of Commons Paper).
5. K. F. Glover and D. N. Miller, 'The Outline of the Road Goods Transport Industry'. (*Journal of the Royal Statistical Society*, Series A, Vol. 117 (1954), pp. 297–330). Report of 1952 Sample Survey.
6. *The Transport of Goods by Road-Report of a Sample Survey made in April*, 1958. (Ministry of Transport, 1959).
7. K. F. Glover, ' Statistics of the Transport of Goods by Road '. (*Journal of the Royal Statistical Society*, Series A, Vol. 123, (1960), pp. 107–139). Gives further details of the 1958 Survey, and relates them to other figures.
8. *Census of Production Reports* (Board of Trade). Annual.
9. Census of Distribution. (Board of Trade). *Census of Distribution and Other Services*, 1950. (Particularly Vol. III. Wholesale Trades (1955)). *Report on the Census of Distribution and Other Services*, 1957 (Board of Trade, 1959).

10. *Road Fund.* Report on the Administration. (Ministry of Transport). Up to 1955–6.
11. *Roads in England and Wales. Report for the Year.* (House of Commons Paper). Annual from 1956–7.
12. *Industry and Employment in Scotland and Scottish Roads Report.* (Command Paper). Annual from 1956–7.
13. *Road Research—The Report of the Road Research Board with the Report of the Director of Road Research* (Department of Scientific and Industrial Research). Annual.
14. W. H. Glanville and R. J. Smeed, ' The Basic Requirements for the Roads of Great Britain ' (Institution of Civil Engineers, 1958. From *Proceedings of the Conference on the Highway Needs of Great Britain,* 1957).
15. *Annual Bulletin of Transport Statistics for Europe* (United Nations Economic Commission for Europe).
16. *Economic Trends* (Feb. 1960, no. 76). ' Statistics of Inland Goods Transport.'
17. J. C. Tanner. ' The Sampling of Road Traffic '. (*Applied Statistics,* Nov. 1957, pp. 161–170).
18. *Statistics of Public Road Passenger Transport in Great Britain* (Ministry of Transport). Annual. (This began with figures for 1949 and 1950. Formerly the information was published in [19]).
19. *Summary of Annual Reports of the Traffic Commissioners* (Ministry of Transport). Annual. (Formerly called *Licensing Authorities for Public Service Vehicles* from the end of World War II to 1955–6. In 1956–7 the name reverted to the pre-World War II title).